REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 12-19-89	3. REPORT TYPE AND DATES COVERED Final Scientific Report 6/86 - 6/89

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
(U) Quantum-Theoretical Methods and Studies Relating to Properties of Materials	

6. AUTHOR(S)

Carl S. Ewig

DTIC
ELECTE
FEB 08 1990
S D

61102F 2303/A3

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Vanderbilt University Nashville, Tennessee	AFOSR-TR. 90-0126

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR/NC BOLLING AFB, WASHINGTON, DC 20332-6448	Grant AFOSR-86-0146

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for Public Release: Distribution is Unlimited	

13. ABSTRACT (Maximum 200 words)

This Final Scientific Report summarizes work completed under research grant AFOSR-86-0146 for the project entitled Quantum-Theoretical Methods and Studies Relating to Properties of Materials. This research concerned the development of new ab initio nonempirical quantum-theoretical methods and computational techniques for studying molecular properties related to those of advanced materials. Extensive computations were performed to demonstrate the quality of the theoretically predicted results. These techniques were employed to study in detail several specific molecular species, focussing on those with unusual and potentially useful energies, structures, spectra, and related properties.

14. SUBJECT TERMS		15. NUMBER OF PAGES 33
Ab initio; Quantum Theory	Molecular Spectra Properties of Materials Computational Chemistry	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL (UNLIMITED) SAR (SAME AS REPORT

NSN 7540-01-280-5500

90 02 06 233

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

FINAL SCIENTIFIC REPORT OF GRANT AFOSR-86-0146

Covering the Period from June 1, 1986 Through June 30, 1989

for the Project entitled

Quantum–Theoretical Methods and Studies
Relating to Properties of Materials

Submitted to the

Air Force Office of Scientific Research, AFSC

by

Carl S. Ewig
Department of Chemistry
Vanderbilt University
Nashville, Tennessee 37235

Table of Contents

Accesion For

NTIS CRA&I
DTIC TAB
Unannounced
Justification

By
Distribution

Availability Codes

Dist   Avail and / or
      Special

A-1

i

FINAL SCIENTIFIC REPORT FOR GRANT AFOSR-86-016
Submitted by Carl S. Ewig, Vanderbilt University

Abstract

This Final Scientific Report summarizes work completed under research grant AFOSR-86-0146 for the project entitled Quantum-Theoretical Methods and Studies Relating to Properties of Materials. This research concerned the development of new *ab initio* nonempirical quantum-theoretical methods and computational techniques for studying molecular properties related to those of advanced materials. Extensive computations were performed to demonstrate the quality of the theoretically predicted results. Finally, these techniques were employed to study in detail several specific molecular species, focussing on those with unusual and potentially useful energies, structures, spectra, and related properties.

I. Introduction

The pressing national need to develop advanced new materials with specific structures, energies, and similar properties of practical importance has led to a variety of theoretical approaches, both empirical and nonempirical, to aid in this development. One of the most promising of these is the use of nonempirical *ab initio* quantum-theoretical techniques, which has been made practicable by the most recent advances in computer architecture, as well as in semiconductor design and software development. Although for many years the size of the species to be studied has precluded rigorous theoretical study of nearly all systems of practical interest, during approximately the past five years there have been very rapid advances in applications of quantum theory to molecules, solids and surfaces that are now beginning to revolutionize aspects of materials development.

This Scientific Report summarizes work long several related approaches that are closely related to the use of nonempirical theoretical procedures for understanding and prediction of unique molecules with specific and potentially useful characteristics and properties. This work included studies involving purely theoretical developments, some tests of various computation models and approximations, and computational investigation of individual molecular species that are of interest in their own right or were intended to illustrate novel types of chemical bonding and related characteristics such as concomitant

thermodynamic quantities. The project was supported in full or in part by Grant AFOSR-86-0146 throughout the period June 1, 1986 to June 30, 1989.

The following report first summarizes our computing resources that proved to be essential to the successful completion of a project of this type. We then discuss the purely theoretical aspects of the project. In particular, we focus on the theory of molecular energy partitioning, which we found to be a useful way of studying the energetics and structural stabilities of a range of molecules and ions, as described subsequently. In the next section we present results of our studies of the accurate computation of properties, specifically (1) basis set and geometry effects on vibrational frequencies, (2) the "gauge-invariant" technique which permits meaningful theoretical prediction of ^{29}Si nuclear-magnetic resonance chemical shifts (a spectroscopy central to the synthesis of silicon compounds), and finally (3) an analysis of procedures for *ab initio* computation of thermodynamic quantities, particularly the enthalpy ΔH°_{298}, in a wide range of species but especially those containing carbon.

The next section of this report then presents the results of three studies of individual compounds or closely related families of compounds that are of interest for some particular aspect of their structures or energies. The first (1) concerned the structures and energetics of P–O–P bridged compounds, focussing on the structural parameters and thermodynamics that characterize this "high-energy phosphate bond" that is ubiquitous in glasses and other solids containing phosphorus. The second (2) focussed on tartaric acid and its low-molecular-weight esters, a prototype of complex species with several conformational degrees of freedom for which accurate experimental vibrational data are available to confirm theoretical predictions. A third study dealt with (3) the structures and vibrational spectra of dibromoethanes, which not only are an additional example of a key species with good experimental vibrational spectra for comparison, but also comprised a useful test of current *ab initio* methods applied to compounds containing elements from the third row of the periodic table.

One of the most significant conclusions to emerge from this project was the discovery that many commonly held concepts (based on laboratory synthesis or early theoretical work) of the instability chemical bonding of the numbers of atoms or groups connected to a common center are often inapplicable. In other words quantum theory unambiguously predicts that several such species with unusually high degrees of bonding to a common atom are structurally stable, even though the apparent difficulty in their synthetic preparation has thus far prevented them from being observed experimentally. We summarize below three classes of these rather remarkable species. The first are (1) the hydrides of second-row elements with the formulas PH_n, SH_n, and ClH_n where n may be as high as five for PH_n, six for SH_n and five for ClH_n. In particular this work was the first indication of the possible existence of the higher chlorine hydrides, ClH_3 and ClH_5. We subsequently studied (2) the first-row elements with covalent bonding to more than four substituents. We were able to demonstrate theoretically the structural stability of the

pentacoordinate nitrogen compounds NF_nH_{5-n} for n = three to five. This was the first firm evidence for the possible existence of a first-row element with stable covalent five-fold bonding. Examination of the properties of the higher fluorides, particularly the fluoride affinity of NF_5, strongly suggested the stabilities of even higher degrees of coordination. We found, employing the same theoretical procedures, that (3) the six-fold bonded species CF_6^{2-}, NF_6^- and OF_6 should be structurally stable. NF_6^- is thermodynamically stable as well with respect to essentially all likely decomposition products *in vacuo*. Our work on OF_6 is incomplete, but a detailed analysis of the carbon and nitrogen hexacoordinated anions was carried out in terms of individual bond strengths and overall electronic structures. As a referee pointed out when this work was submitted for publication, these results mean that "a great deal of chemistry will have to be re-thought."

The final topic of investigation to be described in this report concerns the use of several quantum-theoretical techniques and models to study some unusual electronic states of common species forming impurity centers in solids. These are excited electronic states of anions in ionic lattices which, although they are central to the spectroscopy and energy flow at or near the surfaces ionic crystals, had not previously been characterized theoretically. They are unique in that they are states that cannot exist *in vacuo* since they are unstable with respect to electron autodetachment. They owe their existence entirely to interaction with the (possibly truncated) Madelung potential of the surrounding crystalline solid. They hence have no analog in any other type of molecular species. Our first study concerned (1) developing an accurate theoretical model of the impurity center electronic states and their coupling with the ionic environment. Focussing on the experimentally well-characterized CN^- ion, we were able to demonstrate conclusively that the previously unassigned emission spectrum of alkali halide crystals excited by synchrotron radiation is due to the formation of a lattice-stabilized $^3\Sigma^+$ state of CN^-. In order to check and refine our theoretical techniques we subsequently examined (2) the ground state of CN^- in some detail, including the extent of the ionic lattice included in the computation, the role of the lattice constant determined by the composition of the alkali-halide host, and the perturbation of the impurity species (in terms of orientations, bond distance and vibrational spectrum) due to the presence of the electron distribution on the host ions. We showed that for the special case of the UV/visible spectroscopic properties of the impurity center, the results are remarkably insensitive to precise way in which the lattice is modeled is represented, although the electrostatic potential is in each case essential to existence of the excited states by stabilizing against electron autodetachment.

In the following sections of this report we describe our findings in each of these areas.

II. Computation Techniques and Resources

Since some parts of this research program required extensive computations, our groups' computational resources that were developed during the course of this grant will be summarized separately. The VAX 11/750 system, acquired under Grant AFOSR-85-0072, has now been running continuously without significant failure since its installation in April, 1985. In late 1987 we participated in a proposal by the Department of Chemistry to the National Science Foundation under the Chemistry Instrumentation Program for a SCS40/4 minisupercomputer. The proposal was funded in full. There was also a major contribution from the manufacturer and Vanderbilt University. The NSF award also included three years of maintenance. This new system was installed early in April, 1988, has been running continuously and with almost no repairs or unscheduled maintenance since its installation. The SCS40 is a "Cray clone" which emulates a Cray X/MP computer. It therefore makes use of the vector architecture of the Cray. It has a single processor with four megawords of memory. It quantum-chemical applications it consistently runs at 30% the speed of a Cray X/MP48 (with four processors and eight megawords). We are a primary user of this new system, and it enormously assisted in our research during the last year of this grant.

Acquisition of the new SCS system necessitated implementation of new software libraries designed specifically to take advantage of the system's vector architecture. Therefore as part of this research project we implemented five major program packages on the SCS. These consist of both multi-function program systems, and more specialized programs tailored for specific applications. All have been suitably modified for our computing environment, and several are advanced "beta test" versions available only at a few other laboratories.

Much of our computation is still done by GAUSSIAN86. This is a large multi-function program system for *ab initio* computations that has several technical advantages over the earlier version, GAUSSIAN82, that are important for our work, particularly faster algorithms for electron-correlation energies and more efficient use of disk space in analytical frequency calculations.

For energies employing second-order perturbation theory (MP2) for the electron-correlation effects we have a specialized version of HONDO that was constructed in our laboratory for this purpose, which is about three times faster than the GAUSSIAN programs. However this has now largely been supplanted by a new program system, CADPAC, which was written by Prof. N. Handy's group at Cambridge University. CADPAC is highly oriented toward MP2 calculations, not only for computing energies but also energy gradients and energy-optimized molecular structures. The latest version includes analytical MP2 vibrational frequencies as well. We acquired this system early in 1988. This was apparently the first complete installation of this advanced CADPAC release in the United

States. The original Cray version, written in highly vectorized code, was subsequently implemented by us on the SCS-40 computer.

As noted above, for highly precise calculations and those involving excited or paramagnetic electronic states we employ the UEXP2 (unitary-group exponential-operator) MCSCF program kindly provided by Dr. Ron Shepard at Argonne National Laboratory. Although the input format for this program can be daunting, we have found that it is extremely efficient and reliable, able to treat typically 5000 optimized configuration state functions with little difficulty on either the VAX computer in our laboratory or in the University Computer Center. It has also recently been implemented on the SCS-40, where it is now being used extensively.

There are presently two versions of the Comenius program, specifically for employing coupled-cluster and high-order perturbation methods for the electron-correlation energy, that are presently running on our computers. The first is the efficient code for CCD calculations (the coupled-cluster expansion beginning with double excitations) using molecular symmetry, which was written in our laboratory by Dr. Petr Cársky about two years ago. The second was prepared by Dr. Vladimir Kellö while working in our laboratory during the summer of 1988. It includes, in addition to CCD, the CCSD and CCSDTn (n up to 3) methods plus many-body-perturbation theory through fourth order. It will soon be extended to the full CCSDT level (employing the complete coupled-cluster expansion expansion based on single, double and triple excitations). When operational it will be the second installation of such a program in this country.

Also during the summer of 1988 we implemented the GAMESS system on our SCS-40 computer. This is a modern multi-function program system which contains some particularly powerful and efficient programming for energies and energy gradients in the configuration-interaction (CI) and multiconfiguration self-consistent-field (MCSCF) formalisms using modern unitary-group techniques. It contains useful algorithms for analytical MCSCF gradients and hence molecular structures and vibrational frequencies. It is proving to be an especially useful way of studying the energies and other properties of excited electronic states. We are now using the February 1989 version (with newer code for some sections) and are working with its authors in its continued development.

Table I.1. below summarizes the uses of the primary multi-function program systems for molecules and clusters. All of these program systems also can compute equilibrium molecular structures and the moments of the electronic charge distribution. The following table, Table I.1., summarizes the capabilities of program systems which specialize in precise energy computations for molecules and clusters.

Table II.1. Capabilities of Multi-Function Program Systems

program	applications
GAUSSIAN 86	Energies (SCF, MP2 to 4, CID&CISD, CCD, GVB-PP and some subclasses), analytic vibrational frequencies (SCF only), polarizabilities & hyperpolarizabilities (SCF), infrared and raman intensities (SCF&MP2)
CADPAC Issue 4.0	Energies (SCF,MP2,MP3), analytic vibrational frequencies (SCF&MP2), polarizabilities and hyperpolarizabilities (SCF&MP2), infrared intensities (SCF&MP2), raman intensities (SCF), quadrupole derivatives, frequency-dependent polarizabilities, magnetizabilities (SCF)
GAMESS	Energies (SCF,GVB,MCSCF,CI and all subclasses), localized-orbital analysis, transition moments between electronic states, intrinsic reaction coordinates, analytic MCSCF and CI energy gradients with geometry optimization & vibrational frequencies, energies & other properties of excited electronic states

Table II.2. Capabilities of Programs for Molecular Energies

program	methods
UEXP2	MCSCF (unitary-group exponential-operator approach)
COMENIUS	Many-body perturbation theory (orders 2 thru 4), coupled-cluster expansions CCSD, CCSD+T(CCSD),CCSDT1, CCSDT2 and CCSDT3

III. New Quantum-Theoretical Methods

A. Perturbative Corrections to Basis Incompleteness

There were two major quantum-theoretical efforts during the course of the research project. The first concerns employing perturbation theory to correct the incompleteness of the basis sets used to describe the orbitals in *ab initio* computations. The limited size of the set of basis functions needed to describe accurately molecular electron distributions is a major problem in nearly every application of quantum-chemical computations, since practical considerations of computer time and memory require that the set must be truncated. However it was pointed out several years ago by McDowell[1] that the effect of this truncation, that is the incompleteness of the set, could be approached by means of perturbation theory. More explicitly, if additional basis functions not in the original set are deemed necessary for accurate results, their effect on the energy may be found via a perturbation of the wavefunction (and energy) obtained with the smaller set, without repeating the whole computation.

Dr. Lidia Smentek-Mielczarek, from the Copernicus Institute in Torun, Poland, was visiting our research group during the summer of 1987. Dr. Smentek-Mielczarek is an expert in the field of perturbation theory, and studied the basis-incompleteness problem carefully. She succeeded in generalizing McDowell's formulation, which is based on a Green's function approach to the perturbative solution of the Dyson equation. We then jointly formulated a computer program embodying the mathematics for computing molecular energies and wavefunctions at the SCF level of approximation. This was based on a program originally written in McDowell's research group. We then carried out a series of test computations to determine the method's efficiency and the accuracies of its several levels of approximation. In particular we studied for the first time the adequacy of the perturbation approach to basis sets employed in modern quantum chemistry, such as 6-31G, 6-31G* and 6-31G**.

As an example Table III.1 below show the perturbative corrections to the computed energies of H_2O employing the split-valence 6-31G basis by completing it with d-symmetry polarization functions on the oxygen atom. This is denoted as 6-31G →6-31G* in our notation. The table also shows the 6-31G →6-31G**, which adds both d functions to the oxygen and p functions to the hydrogens, roughly doubling the size of the basis. The rows in this table represent various levels of approximation obtained by omitting some of the more time-consuming terms in the complete perturbed energy expression. This also shows for comparison the exact energy difference obtained from a full conventional calculation of the energy in each basis. Clearly the perturbation approach, especially approximation V, is exceptionally accurate. Even in the case of the 6-31G →6-31G** perturbation the remaining error is only 0.000034 hartrees or 0.089 kJ/mol.

Table III.2 shows a similar comparison for the effect of the basis set on optimized molecular structures. The energy in each case was varied in a pointwise fashion to obtain

the equilibrium structural parameters. This was the first such comparison employing the basis-set perturbation approach.

Table III.1.   Corrections to molecular energy[a] of H_2O from the basis-incompleteness perturbations 6-31G→6-31G* and 6-31G→6-31G**.

approximation	6-31G→ 6-31G*	6-31G→ 6-31G**
I	−.020938	−.034474
II	−.021071	−.034892
III	−.023927	−.036540
IV	−.024017	−.036778
V	−.024327	−.037238
exact difference between basis sets	−.024322	−.037272

[a]Energy in the larger basis set minus the energy in the smaller basis in hartrees at the 6-31G optimized molecular structure. The total energy in the 6-31G basis, to which these values are to be added, is −75.985359 hartrees.

Table III.2.   Effect on computed equilibrium structure[a] of H_2O from the basis-incompleteness perturbations 6-31G→6-31G* and 6-31G→6-31G**.

approximation	6-31G→6-31G*		6-31G→6-31G**	
	Δ OH length	Δ HOH angle	Δ OH length	Δ HOH angle
I	−0.00084	−5.12	−0.00453	−4.69
II	−0.00080	−5.13	−0.00479	−4.76
III	−0.00242	−5.95	−0.00630	−5.38
IV	−0.00236	−5.94	−0.00626	−5.36
V	−0.00230	−6.06	−0.00653	−5.57
exact difference between basis sets	−0.00232	−6.04	−0.00658	−5.58

[a] Bond lengths and angles computed in larger basis minus those in the smaller basis. Lengths in Å, angles in degrees. The values in the 6-31G basis, to which these values are to be added, are OH = 0.94963 Å and HOH = 111.55°.

Again the agreement with the conventional is excellent, especially at the highest level of approximation.

During the summer of 1988 Dr. Smentek-Mielczarek returned to our group and studied the problem of extending this perturbation theory to beyond-SCF approximations, that is, including approximations to the electron-correlation energy. In this formulation there are two perturbations at once, due to basis set incompleteness and also to the correlation. From a computational point of view this would be especially appealing since the same integrals required for the conventional Møller-Plesset many-body-perturbation theory of electron-correlation effects are also generated during the basis-incompleteness perturbation calculation at the SCF level. A theory through second order in electron correlation was developed, and a test computer program was written and tested. However it was found that all of the the required terms cannot be derived in this way, and that computations omitting these terms are of insufficient accuracy. Therefore this line of investigation has not been pursued.

A detailed account of the mathematics required by this approach at the SCF level of approximation, as well as a summary of all our test computations, has been published.[2]

B. Partitioning Theory of Ab Initio Molecular Energies

A second area of primarily theoretical research under this grant concerned the interpretation of *ab initio* computations, rather than the mechanics of how the computations are carried out. In the interpretation of the general features of molecular structures and energetics, as studied on either an experimental or theoretical level, there has historically been recourse to descriptions invoking the interactions between specific individual atoms or sets of atoms. Many commonly accepted rationales invoke the presence of strong bonding interactions between specific atoms, interactions through space between atoms not considered to be bonded, as well as multicenter bonding, high-energy bonds, *etc.* However to our knowledge no general theoretical approach has been developed for testing a given molecular electronic structure to determine the pattern and extent with which the energy is distributed among atoms.

Since it is natural in the analysis of computed results, such as the stabilities of the hypercoordinated species as described below in Section VI of this report, to employ this picture based on interactions among atoms, we developed a complete theory and associated set of computer programs for the partitioning of molecular energies in terms of energies associated with a given nucleus only, specific pairs of nuclei, triplets of nuclei, *etc.* This is to some extent arbitrary, since each basis function assumes nonzero values throughout the molecule (except on nodal surfaces). However, as has been pointed out,[3] it is no more or less arbitrary than discussing molecules generally in terms of their constituent atoms.

Thus the basic principle is that the energy of any system, with a specific geometric and electronic structure, may be expanded in terms of individual energies associated with centers or sets of centers (generally atomic nuclei),

$$E^{Total} = \sum_n E_n^{(1)} + \sum_{mn} E_{mn}^{(2)} + \sum_{lmn} E_{lmn}^{(3)} + \dots \qquad (1)$$

where the summations are over all centers in the system. The $E_n^{(1)}$ will be referred to as one-center energies, the $E_{mn}^{(2)}$ as two-center energies, *etc.*

We first showed, using the second-quantization formalism, that for any *ab initio* wavefunction employing basis functions centered on nuclei (which is almost invariably the case) the total energy of any species may be rigorously expressed by an equation of the form of (1) above in which there are no more than four such summations, *i.e.* multiplicities of atoms higher than quartets cannot simultaneously appear in the energy. We derived general expressions for these energies in terms of the one- and two-electron density matrices. In each case, once the density matrices have been determined, the $E_n^{(1)}$, $E_{mn}^{(2)}$), *etc.* may be written in closed form.

For the special case of the SCF approximation we showed how the general results reduce to algebraic expressions for each of the four types of energy terms. For example the one-center energy of atom n, $E_n^{(1)}$, is given by the sum of kinetic, electron-nuclear attraction and electron-electron repulsion terms,

$$E_n^{(1)} = \sum_{ij \text{ on } n} d_{ij} (\langle i|-\nabla^2/2|j\rangle + \langle i|-Z_n/r|j\rangle) + 1/4 \sum_{ijkl \text{ on } n} g_{ijkl} (2d_{ij}d_{kl} - d_{ik}d_{jl}).$$

where d_{ij} are elements of the density matrix over basis functions (the "bond-order matrix"), Z_n is the nuclear charge, r the distance from the nucleus, and g_{ijkl} are the two-electron repulsion integrals. We then showed the results of sample computations of several properties of molecular energies, including *(1)* a detailed resolution of all the terms appearing in the molecular energy of H_2O. This is shown below in Table III.3 since it provides a good example of the type of values that are obtained. Note that, not surprisingly, the total energy is dominated by the $E^{(1)}$ term of oxygen. The two-center H–O energies are negative and hence stabilizing. There is a repulsion between the two hydrogens, and a modest H–O–H three-center stabilization.

In addition we also reported (2) the fairly small basis-set dependence of the one- and two-center energies for the diatomic species H_2, Li_2, N_2, and F_2, *(3)* the convergence of the energy expansion upon going from $E^{(1)}$ to $E^{(2)}$, $E^{(3)}$, *etc.* in H_2O, N_2O, CO_2, NH_3, and NF_3, *(4)* the one-, two- and three-center contributions to the energy of the reaction $2H + O \rightarrow H_2O$, *(5)* a similar analysis for three reactions among diatomic molecules, $H_2 + F_2 \rightarrow 2HF$, $Li_2 + F_2 \rightarrow 2LiF$, and $H_2 + Li_2 \rightarrow 2LiH$, *(6)* the values of the derivatives $\partial E^{(1)}/\partial R$, $\partial E^{(2)}/\partial R$, $\partial^2 E^{(1)}/\partial R^2$, and $\partial^2 E^{(2)}/\partial R^2$ for a set of ten diatomic species, where R is the internuclear distance, and *(7)* a detailed analysis of the individual energy terms contributing to the derivatives $\partial E^{(1)}/\partial R$ and $\partial E^{(1)}/\partial \alpha$ for H_2O, where R is the symmetric O–H stretching coordinate and α is the H–O–H angle. Except for the tests of the basis-set

10

Table III.3. Detailed Multicenter Resolution of the SCF Energy[a] of H_2O

One-Center Terms

Center	Kinetic	E/N Attraction	E/E Repulsion	Total
H_1	0.37299	−0.55737	0.03821	−0.14617
H_2	0.37299	−0.55737	0.03821	−0.14617
O	74.23529	−176.46529	28.05106	−74.17894
Total	74.98126	−177.58003	28.12749	−74.47128

Two-Center Terms

Centers	Kinetic	E/N Attraction	E/E Repulsion	N/N Repulsion	Total
H_1 O	0.51636	−10.55520	4.77268	4.41902	−0.84714
H_2 O	0.51636	−10.55520	4.77268	4.41902	−0.84714
H_1 H_2	−0.00295	−0.23974	0.04992	0.34930	0.15653
Total	1.02976	−21.35013	9.59528	9.18733	−1.53776

Three-Center Terms

Centers	E/N Attraction	E/E Repulsion	Total
H_1 O H_2	−0.22049	0.18133	−0.03916
Total	−0.22049	0.18133	−0.03916

[a]In hartrees, computed at the experimental equilibrium structure, in the 6-311G(2d,p) basis. Total SCF Energy = −76.04820.

effects, these were all carried out in the fairly large 6-311G(2d,p) basis, consisting of a triple-zeta set plus two d functions on the heavy atoms and one p on each hydrogen. The overall conclusion to be drawn from these results is that conventional concepts of molecular structure and energetics must be treated with caution, since most processes, such as lengthening or breaking a chemical bond, generally involve changes throughout the molecule, rather than being confined to a localized two-center contribution. In particular, one-center energies are often dominant. This is of course in accord with modern electronic-structure principles, since the energy is known to be particularly sensitive to the behavior of the electron distribution close to the nuclei, which contributes only to $E^{(1)}$.

Although the above results were obtained at the SCF level, the theory is completely general within the framework of molecular orbitals spanned by atomic basis functions. As an illustration we carried out the reduction of the general expression to the case of second-

11

order Møller-Plesset perturbation theory to include the electron-correlation energy. This is commonly called the MP2 energy. A separate computer program was constructed for this purpose, and some representative test calculations were carried out. For example Table III.4 below shows the multicenter partitioning of the MP2 energy of H_2O. These values

Table III.4. Resolution of Second-Order Perturbation (MP2) Approximation to the Correlation Energy[a] in H_2O.

$E^{(1)}_H$	–0.02435
$E^{(1)}_O$	–0.30087
$E^{(2)}_{OH}$	0.06524
$E^{(2)}_{HH}$	–0.00021
$E^{(3)}$	–0.01448
E^{Total}	–0.23378

[a]In hartrees, computed at the experimental equilibrium structure in the 6-311G(2d,p) basis. Excitations from the 1s orbital of oxygen not included.

may be compared directly with those in Table III.3. The one-center terms are seen to dominate the correlation energy and have values roughly comparable to those in the free atoms (zero for hydrogen). But perhaps the most interesting result is that the "bonding" correlation-energy terms between O and H are positive and hence destabilizing. This is in accord with the usual finding that including electron correlation corrections tends to lengthen bonds and decrease stretching frequencies. We will return to this point in Section IV.

In summary, we believe that the general theory of multicenter partitioning of *ab initio* molecular energies is emerging as an extremely powerful tool for understanding molecular structures, energies, and related properties. We will discuss in particular the analysis of the structural stabilities of unusual and high-energy species below in Section VI.

This work has been submitted for publication. A preprint is attached to this report as Appendix A.

IV. Computed Properties of Materials

A. Basis and Geometry Effects on Vibrational Frequencies.

One of the objectives of our overall research program was to develop and/or test *ab initio* procedures for finding derived properties, such as spectroscopic parameters. In the

12

first such study we examined the role of the choice of basis set and molecular structure in computing vibrational spectra.

The use of specialized basis set to minimize superposition error in computed molecular energies as atoms or groups of atoms are brought together has been proposed by Tatewaki and Huzinaga.[4] Hence it appeared possible that these basis sets might well be advantageous for finding vibrational frequencies, particularly for bond-stretching modes. Since ethylene represents a species whose internal force field has been extremely well characterized, we computed the equilibrium molecular structures, vibrational frequencies and infrared intensities of ethylene in nine basis sets. These included three Pople-type sets as well as six sets of comparable size recommended by Tatewaki and Huzinaga. However it was found that the latter six gave inferior results in nearly every category when compared to Pople-type set of comparable size.

A second question addressed in this work is whether greater accuracy is achieved in computing vibrational frequencies when the molecular geometry is taken to be the computed equilibrium structure, or when the frequencies are derived from the hessian matrix computed at some other geometry such as the experimental equilibrium structure. (The latter can readily account for the Eckhart conditions especially when working in internal coordinates.) It is well known that computations at the SCF level of approximation generally give bond lengths that are too short and vibrational frequencies too high. This is also true to lesser extends for many approximations including correlation effects. Therefore we investigated *(1)* whether stretching frequencies are more accurately computed at experimental rather than computed structures, employing SCF results in the 6-311G** basis for ten representative diatomic species (H_2, LiH, HF, LiF, Li_2, BH, BF, CO, N_2, and F_2) for which accurate experimental vibrational frequencies and internuclear distances are known. The SCF bond length is shorter than the experimental for each species except Li_2 and LiH. For each species except these two the frequency was more accurate at the experimental internuclear distance. A similar comparison was carried out for ethylene, showing that in this case all the frequencies were decreased and brought closer to experiment by employing the experimental geometry. There are similar but smaller improvements in infrared intensities. We also considered *(2)* whether longer internuclear distances are necessarily associated with lower frequencies by analysis of the Morse potentials for the common ranges of their parameters, and within the approximation that the correlation energy (difference between the SCF energy curve and the experimental one) varies linearly with nuclear displacements.

Overall our conclusions reinforced the view that vibrational spectra computed at experimental (or higher-level theoretical) geometries almost always more accurate than those at commensurate computed equilibrium structures. This is reinforced by theoretical analysis of the shapes of molecular potential energy curves. However there are exceptions, such as for the Li species, and the solution of the vibration secular equations to obtain the

frequencies is somewhat complicated (due to Eckhart conditions) so it does not seem advisable to employ experimental geometries routinely.

An account of this work was published early in the project period. See Section IX.

B. ^{29}Si Nuclear-Magnetic Resonance Chemical Shifts

The chemistry of silicon is critical to the design of high-performance materials. For many years ^{29}Si nuclear-magnetic resonance (NMR) was not employed as widely in the preparative chemistry of silicon as other nuclei (such as ^{1}H, ^{13}C, or ^{31}P) due to its low sensitivity, low geomagnetic ratio and relatively low isotopic abundance. However with the recent development of Fourier-transform techniques, it is rapidly becoming a standard tool in both organic and inorganic silicon chemistry. This has led in turn to a need to be able to predict in advance what the NMR spectrum of a new compound will be, particularly in cases where the spectrum of no similar species has been assigned.

This is a field in which *ab initio* quantum methods will play an increasingly important role, since it is now well established that SCF procedures which properly account for the gauge of the magnetic field can give NMR spectra with chemical shifts that are of impressive accuracy when compared to the uncertainties in assigning experimental spectra. Therefore we carried out the first comprehensive study of a range of silicon compounds employing the gauge-invariant atomic orbital (GIAO) approach[5] to compute ^{29}Si chemical shifts. This work followed from Section III.F of our original research proposal.

Alltogether ^{29}Si nmr shielding tensors (and the chemical shifts derived therefrom) were calculated for 28 representative silicon compounds, employing optimized molecular structures and a moderately sized basis set, both with and without d functions, with the gauge-invariant GIAO approach. Differences in shielding between members of the following substituent-substitution series of compounds were investigated: SiH_4/SiF_4, $SiH_4/Si(CH_3)_4$, $Si(CH_3)_4/SiF_4$, and part of two series involving the $SiCl_4$ molecule. Since these series of molecules show a wide range of ^{29}Si chemical shifts, with the shift values plotted for some series exhibiting a pronounced hump rather than a linear form, they represent a good test of the theoretical approach employed. Within these series and within a group of silyl derivatives, the calculated and experimental chemical-shift data agreed quite well. A group of disilicon compounds, H_3Si-X-SiH_3 with X = nothing, O, NH, and CH_2, were also studied, as was the SiF_6^{2-} anion and some other structures considered to be exemplificative. The effects on the shielding of rotational isomerization as well as of induced geometrical and electronic changes were investigated for several molecules. The paramagnetic and diamagnetic contributions to the magnetic shielding were analyzed, and an apparent relationship between the ^{29}Si paramagnetic term and the electron-withdrawing power (EWP) of the substituents on the silicon was found. An approximate relationship between the paramagnetic and diamagnetic terms, via the EWP, was also shown to apply to

our computations. These allow a simple interpolation to characterize the paramagnetic shielding in the spectra of unknown compounds based on available data.

We also investigated several technical details of these calculations, including a comparison with other theoretical approaches (such as the invariant-gauge localized orbitals, IGLO[6]), the importance of employing experimental versus computed equilibrium molecular structures, the size of the basis set (including the role of d functions on the silicon atom), and a detailed analysis of the effect of basis set size on individual terms that contribute to both diamagnetic and paramagnetic shielding tensors in SiH_4 and SiF_4.

One of the more interesting findings from this work was that when employing relatively small basis sets, such as 6-31G and 6-31G*, remarkably accurate *differences* in the chemical shifts between differing compounds may be found, although the absolute shieldings may be appreciably in error. Also, these computed differences computed using a gauge-invariant procedure compare very well with other types the require much larger basis sets. This is particularly true of the larger species which are of practical interest. For example Table IV.1 below compares the computed [29]Si chemical shifts for bridged disilyl compounds with experiment.

Table IV.1. Comparison of Calculated and Experimental [29]Si Chemical Shifts (in ppm referenced to TMS) of Some Bridged Disilyl Compounds.

Compound	Experimental	Calculated[a]	
		No d	With d
H_3SiSiH_3	-104.8	-89.5	-78.6
$H_3SiOSiH_3$	6.7	-23.7	-32.0
$H_3SiNHSiH_3$	2.2	-36.8	-37.7
$H_3SiCH_2SiH_3$	-60.9	-49.2	-44.0

[a]The reference is tetramethylsilane with $\sigma_{TMS} = 484.5$ for "no d" and 447.9 for "with d."

To show the efficiency of this procedure for shift differences, even when employing relatively small basis sets, we show below in Table IV.2 a comparison of our gauge-corrected results with two basis sets, 6-31G and 6-31G(*), the latter including d functions on the Si atom. These are compared with a conventional "gauge-variant" computation and with experiment. Even without the d polarization functions the agreement of the shift differences (except for SiO_4^{4-}) with experiment is acceptable. The agreement was found to be much better comparing members of the substitution series replacing one type of atom or group systematically with another going from SiH_4 to SiF_4, SiH_4 to $Si(CH_3)_4$, and $Si(CH_3)_4$ to SiF_4.

Table IV.2. Comparison of the Gauge-Variant and Gauge-Invariant (This Work) Computed Chemical Shifts With Experiment, Referenced to Silane.

Molecule	Gauge-Variant[a]	6-31G[b]	6-31G(*)[b]	Experimental
Si_2H_6	-21	-18	-9	-13
SiF_4	-76	+25	-18	-21
SiF_6^{2-}	-131	-46	-97	-95
SiO_4^{4-}	-31	-40		+20

[a]Referenced in each case to silane. [b] Our GIAO results.

We conclude that the "gauge-invariant" GIAO approach is useful one for describing ^{29}Si NMR, predicting shielding tensors and chemical shifts with surprising accuracy even when employing the smaller types of basis sets that are the practicable ones for treating species of the sizes that are of potential interest to silicon chemists.

Finally, a detailed account of this work has been published.[7]

C. Reaction Enthalpies

The thermodynamics of reacting species is the key determinant in the course of many classes of chemical reactions, especially those involving inorganic compounds. As a part of this research project (proposed in Section III.D of our original research proposal) we began a systematic study of the accuracy with which thermodynamic properties, especially enthalpies, could be obtained using currently available quantum-theoretical procedures. This project consisted of three basic steps. First we constructed a data base containing total molecular energies (electronic and vibrational at 298 K) for a number of differing chemical species, each at a number of differing levels of approximation. The approximations included choice of basis set, correlation-energy procedure (many-body perturbation theory, coupled cluster theory, etc.), choice of molecular structure (experimental or computed) and vibrational energies computed with differing levels of approximation. The choice of species for study initially focussed on those containing second-row elements such as silicon since for these the thermodynamics is particularly important in determining their chemistries. We devoted most of our attention in this initial stage to sulfur since of the second-row elements it has the most extensive experimental data for comparison. However as the work developed it soon became apparent that all nearly the currently available methods were inadequate for yielding enthalpy differences accurately enough to be practically useful, so we widened the scope of our computations to include

first-row elements as well. The resulting data base has continued to grow, and now contains 1347 entries.

Clearly a file of this size cannot be manipulated manually, so we developed a suite of computer programs to extract and analyze the values in the data base in terms of enthalpies of reactions. In computed enthalpy differences it is essential to include only values that were obtained at exactly the same level of approximation, thus for example not permitting an energy computed at the MP4 level for reactants but SCF or coupled-cluster theory for products. Our new program takes as input the data base plus any reactions written in conventional chemical notation, parses them in terms of stoichiometries and individual energies at strictly commensurate levels of approximation, and gives the results in tabular form for each reaction. When large numbers of reactions are to be scanned, this input is prepared as a separate file. The output tables are then edited and condensed.

The results of this project are still being organized and written. Of the first three publications, one has thus far been submitted for publication and two more are in preparation. The manuscript of the first is attached to this report as Appendix B. In this work *ab initio* computations were performed to obtain the enthalpies of over 40 gas-phase diamagnetic species including 18 hydrocarbons. These all employed optimized geometries and a wide range of Pople-type basis sets (from 4-31G to 6-311++G(2df,2dp)) and correlation-energy approximations (including Møller-Plesset series from MP2 to MP4SDTQ and coupled-cluster theory from CCD to CCSDT). The electronic energies of forming the molecules from the electrons and bare nuclei were derived from experimental data and compared with the *ab initio* results at the various levels of approximation. Then the enthalpies of chemical reactions at 298 K between series of molecules were computed for several generic classes of reactions, including dehydration of the common hydrides and combination of homonuclear diatomics. Although several hundred chemical reactions were computed, we explicitly listed the results for the representative reactions $N_2 + 3H_2 \rightarrow NH_3$ and $CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$. Finally we examined the enthalpies of isomerization of 7 hydrocarbons containing 3 to 4 carbon atoms.

As a result of this lengthy series of comparisons we were obliged to conclude that present quantum-theoretical procedures, including basis sets and correlation-energy procedures, are at best marginally accurate in describing the enthalpies of simple chemical reactions (within say 1 kcal/mol). This applies even to the highest levels of available procedures (largest basis set and highest-order perturbation approximation to the energy). We found that the basis set plays central role in determining the accuracy of the enthalpy, particularly when electronegative elements such as oxygen are involved. This is remarkable in view of the effort that theoretical chemists and physicists have consistently devoted to more elaborate correlation-energy procedures, rather than to developing basis sets. Diffuse Rydberg-type functions are essential in many cases to obtaining even qualitative accuracy. An exception to these generalizations are reactions strictly among the

hydrocarbons, which were less sensitive to the choice of basis set. Hydrocarbon enthalpies were found to be very satisfactorily computed with the CCD approximation.

We were assisted in much of this work by Dr. Vladimir Kellö from Comenius University in Bratislava, Czechoslovakia, who worked in our group during the summer of 1988. Dr. Kellö is a leading expert in the coupled-cluster theory, and also has published extensively in the field of *ab initio* reaction enthalpy computations. The higher-level coupled-cluster methods we employed, such as CCSDT3, are not yet widely used but were made possible by Kellö's implementation on our computer systems of the COMENIUS computer program (summarized above in section II). Subsequent work that will soon be published in this field will include explicit comparisons of the relative efficiencies and accuracies of several variations of the coupled-cluster formalism.

As noted above, systematic examination of many different chemical reactions in this way shows that the hydrocarbons are unique in several ways. In particular for some key reactions the CCD approach gives remarkably accurate results, often within experimental error. However the theoretical reasons for this are not apparent. A summary of all our computed results for hydrocarbons, including relatively high-level computations on species with up to four carbon atoms, will be the subject of a subsequent publication.

V. Studies of Individual Representative Species

In this section we summarize studies that focussed on specific compounds or closely related families of compounds considered to be exemplificative, particularly with respect to molecular conformational preference, relative energies, and vibrational spectra that characterize them. These are in response to Sections III.A-C of our original research proposal.

A. Structures and Energetics of P–O–P Bridges Compounds

In the chemistry of phosphorus, the P–O–P linkage plays a central role, analogous to the C–C linkage in organic species. It appears in a wide variety of environments and applications, including inorganic and biological systems. In particular it is central to phosphorus-containing solid materials such as glasses. However it is readily hydrolyzed, with an attendant large decrease in enthalpy in condensed phases, giving rise to the well-known concept of the "high-energy phosphate bond." Despite its importance there have been remarkably few theoretical attempts to study its properties.

As part of our research project the detailed molecular structures including the stable equilibrium conformations, molecular geometries and relative energies of four prototype phosphorus compounds with a P–O–P linkage, $H_4P_2O_{2n-1}$ for n=1 to 4, were computed. In addition the dipole moments and the changes in both energy and dipole as a function of the torsional angle about the briding P–O bond were determined, as they are central to characterizing it physically. We paid particular attention to the size, bending force constant,

and barrier to linearity of the P–O–P angle (which is known experimentally to be $180°$ in many crystalline solids) in the $n = 1$ to 3 species. Unlike many other related compounds containing a phosphoryl or carbonyl group, the conformations were found to be only modestly affected by purely electrostatic terms as represented by the overall dipole moment function. Fully optimized molecular structures and resulting total energies were reported in each case. The P–O–P angle is always less than $180°$ *in vacuo*. From the *ab initio* computations we also determined the enthalpies and free energies of the reactions $H_4P_2O_{2n}\text{-}1$ + $H_2O \rightarrow 2H_3PO_n$ for $n = 1$ to 4, i.e.the hydrolysis reactions of μ-oxo-diphosphine, phosphinic anhydride, diphosphonic and diphosphoric acids. For the $n = 1$ case we also checked the role of basis set size and of adding electron correlation at the MP2 level. Our results clearly showed that the energies of these reactions, which supposedly release the energy of the "high-energy phosphate bond" decrease slightly with increasing numbers of substituents (higher values of n) but are close to zero in the gas phase. Hence from comparison with experiment the –18 kcal/mol enthalpy change upon breaking the P–O–P linkage often observed in aqueous solutions is due to the molecules' interaction with their environment, particularly hydration of the reaction products, and not a property of the chemical bonding *in vacuo*.

The largest and also the most practically important compound of this series is diphosphoric (or pyrophosphoric) acid. We found two stable conformers of this species. Their computed equilibrium structures are depicted below (drawn to scale).
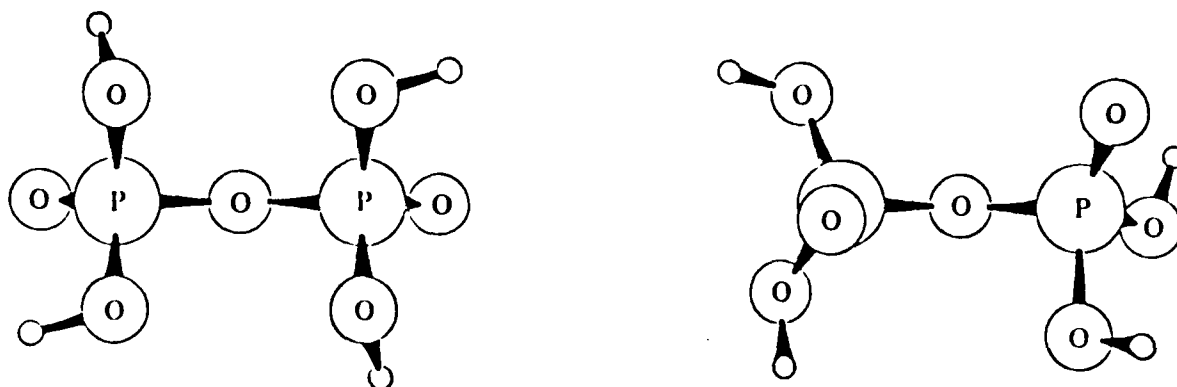
Here the structure on the right represents the more stable conformer. It is interesting that the P–O–P angle is over $8°$ larger in the higher-energy conformer.

From the experimentally determined structures and relative free energies of some of these and related species in condensed phases (particularly in aqueous solutions) it is now clear that they represent an extreme examples of the principle, now becoming widely accepted: that these properties are often determined overwhelmingly by interactions with the surroundings, rather than being properties of the isolated molecules *in vacuo*.

B. Conformations of Tartaric Acid and Its Esters

As the capabilities of our computer programs and other resources increase, there are increasingly the requirements and opportunities to verify the computed results on larger systems. Recent vibrational circular dichroism (VCD) spectra of tartaric acid and some of its esters by Prof. Polavarapu of the Vanderbilt Department of Chemistry provided such an opportunity to study the preferred conformations, their relative energies, and the ways these may affect the observed spectra. Prof. Polavarapu's group studied the spectra of optically active tartaric acid, dimethyl tartrate, diethyltartrate, and diisopropyl tartrate in solution with differing solvents and concentrations. Analysis of the VCD spectra yields information about the preferred conformations. Concentrating on tartaric acid, the structures for the stable equilibrium conformations were computed in the STO-3G and 6-31G basis sets in the SCF approximation. In both cases seven stable conformers were found, corresponding to the various ways in which the hydroxyl and carbonyl groups can internally hydrogen bond. The relative energies of the conformers were then determined at the STO-3G structures in the STO-3G, 3-21G, and 6-31G basis sets, and finally in the 6-31G basis at the 6-31G optimized structure. Significantly, all of the results for the last three approximations were quite similar. Thus, this seems to be a useful way of determining the conformation preferences in relatively large and complex species of this type. All four computations agreed however that conformer with terminal COOH groups oriented trans is lowest in energy. This reinforced the conclusions from the VCD spectra and shed some light on the ways conformation affects observed vibrational spectra in compounds of this type.

C. Structures and Vibrational Spectra of Dibromoethanes

A second set of computations was specifically oriented toward verifying the utility of *ab initio* computations for assigning a set complex vibrational spectra, and also testing the currently available computational procedures. This work was also performed in conjunction with Prof. Polaravapu of this department, who recorded and tentatively assigned the vibrational spectra of 1,2-dibromoethane and the meso and chiral forms of 1,2-dideuterio-1,2-dibromoethane in neat liquid samples and also isolated in low-temperature matrices.

This problem is of particular interest for *ab initio* quantum-chemical study for three reasons: First, it is another case in which there apparently two stable conformers, with Br *trans* or *gauche*, whose relative energies are not conclusively known experimentally but which are known to be much different in the gas phase and in isolated in matrices from other condensed phases. So it was a useful test case to test the behavior of the theoretical computations for the gas phase relative to that in the liquid. Second, we wished to determine the extent that computed nuclear-motion effects, principally due to vibration, influences the relative energies, entropies, and free energies (hence equilibria) among the various conformers. Finally, since bromine is a heavier element than generally encountered in quantum-chemistry, it posed a useful test of whether current computational methods,

especially basis sets, were sufficient for accurately predicting the properties of a species containing a third-row element.

The central computations in this project consisted of finding the *ab initio* molecular structures of *trans* and *gauche* 1,2-dibromoethane, employing both a minimal STO-3G basis and also a "pseudo 6-31G*" basis. The latter was composed of the standard 6-31G* set for carbon and hydrogen, including *d* functions on the carbon, plus the basis of roughly commensurate size for bromine prepared by del Conde *et al.*[8] We then computed the 36 vibrational frequencies of 1,2-dibromoethane, and both the meso and chiral forms of 1,2-dideuterio-1,2-dibromoethane, and finally the infrared intensities of the allowed transitions.

It was found that the available basis sets, particularly the smaller one, are inadequate for describing the molecular structure very accurately. For example for the *trans* isomer the STO-3G basis gives a C-C internuclear distance of 1.545 Å, the 6-31G* basis gives only 1.508, compared to the experimental value of 1.506. For the C-Br distance these distances are 1.916, 1.971 and 1.950 Å respectively. However vibrational frequencies are of reasonable accuracy; for example C-Br stretches are computed to range from 569 to 705 cm^{-1}, compared to the experimental values of 553 to about 660 cm^{-1}. This degree of accuracy in the frequencies, combined with the *ab initio* intensities, was sufficient to assign the complete spectra of all three species. The equilibrium conformation was conclusively shown to exhibit a *trans* orientation of the bromine atoms, as inferred from the gas-phase spectra. Electron correlation at the MP2 level was relatively large, increasing the energy difference from 1.2 to 1.8 kcal/mol. However the nuclear-motion contributions to this free-energy differences were found to be quite small. For example in the undeuterated species the *trans* form has a nuclear-motion (vibration plus rotation) energy E_{298} = 41.78 kcal/mol, compared to 41.61 kcal/mol for the *gauche* form. The entropies are 76.82 and 76.53 cal/(mol K) respectively. Hence the contributions to the relative free energies are also nearly the same. But including it with the other terms increased the energy difference to 2.0 kcal/mol, in excellent agreement with the experimental value in the gas phase of 2.1 kcal/mol.

We concluded from this work that although basis sets for the third-row elements are not suitable for accurate computations of the structures of compounds such as dibromoethane, they are capable of predicting accurate vibrational frequencies, intensities, and conformational preferences.

VI. Unique Highly Coordinated Compounds

Early in this research project we began a systematic study of the possible stabilities of compounds with unusual types of molecular linkages that would be expected to exhibit either very high energies or high thermodynamic stabilities (low energies) relative more

conventional species. There is a broad category of such compounds that have generally been assumed to be "nonexistent" due to the lack of success by synthetic chemists in preparing them. However the quantum-mechanical formulation of possible "existence" had not been developed. Therefore we began a lengthy series of test computations to determine *(1)* could compounds previously dismissed as "nonexistent" could in fact be assigned stabilities on purely theoretical grounds, leading to more serious attempts to prepare them in the laboratory, *(2)* whether currently available methods were reliable and consistent for all the relevant properties (particularly vibrational frequencies and free energies), and *(3)* what are the ultimate features of the molecular electronic structure that contribute to stability, so that generalizations may be made to related compounds.

There are several categories of stability, including electronic, structural, thermodynamic and kinetic. The first is the system's resistance to electron autodetachment. Autodetachment is a problem primarily for small and/or highly charged anions. Structural stability implies that the species will not spontaneously deform or fragment. This is determined by the vibrational frequencies; if all values are real, there is an energy well with positive curvature with respect to all changes in geometry. Thermodynamic stability means that the species has a free energy that is sufficiently low with respect to all likely decomposition products that detectable amounts of the former may exist at a given temperature. Kinetic stability denotes the presence not only of barriers to decomposition, but also sufficiently high barriers that small external perturbations to the species, or tunneling effects, will not permit decomposition in a reasonable length of time. We believe that viewing the problem in this way, and developing the necessary computational tools as described later, is the foundation of the new discipline that might be termed "molecular stability theory."

From a computational point of view the most readily treated is the thermodynamic stability since we now know that free energies may reliably be estimated from *ab initio* computations with at least qualitative accuracy, as discussed in Section IV.C above. But the key property for unusual types of bonding such as described here is the structural stability in terms of the real vibrational frequencies. (From the magnitudes of real frequencies, indirect conclusions may be drawn about kinetic stability, *i.e.* how strongly the atoms are interconnected.) This has been realized for some time, but it has only been in the past four to five years that computer algorithms and programs have become available for computing complete sets of vibrational frequencies (from analytic expressions for the hessian matrix) for polyatomic molecules. Also, only with the advent of supercomputers have the resources become available for studying systems with several heavy atoms such as those described in this section of the report.

A. PH_n, SH_n and ClH_n Species of Maximal *n*

Our first study of this type involved the hydrides of second-row elements phosphorus, sulfur and chlorine with the maximal degree of coordination of the central

22

atom, thus with the formulas PH_n, SH_n, and ClH_n. The species with $n = 3, 2$, and 1 respectively are well known. We wished to determine what other values of n might correspond to structurally and thermodynamically stable species, and what the maximal value of n corresponds to for each element.

It became apparent in the early stages of this project that for this unusual type of bonding the SCF approximation was insufficient, often giving incorrect numbers of real or imaginary vibrational frequencies. Therefore it was determined that electron-correlation corrections with at least the MP2 level of approximation would be needed throughout. Fortunately the most recent version of the CADPAC program package from Dr. N. Handy's group at Cambridge University became available and was installed on our SCS minisupercomputer at this time. CADPAC Issue 4.0 is highly efficient for computation of molecular structures, and is unique in that it computes vibrational frequencies at the MP2 level analytically. To our knowledge it is currently the only program with this capability. The installation of this version in our laboratory was one of the first in the United States. The unique combination of the CADPAC Issue 4.0 program plus the vector-processing capabilities of the SCS computer made possible all of the requisite computations of structural stabilities discussed below.

Five highly coordinated species were found: PH_5, SH_4, SH_6, ClH_3, and ClH_5. The highest possible member of the series, ClH_7, was found to be structurally unstable. This was the first theoretical study of indicating the stabilities of SH_6, ClH_3 or ClH_5. In addition to the optimized molecular structures and the vibrational frequencies, the energies of each species was computed at the SCF, MP2 and MP4 levels of approximation, and in both a 66-31G** basis and for comparison in a relatively large 6-311G(2d,2p) basis.

To assess thermodynamic stabilities we computed the thermodynamic quantities ΔH^o_{298} and ΔG^o_{298} for the hydrogenation reactions (inverse of decomposition by hydrogen loss): $PH_3 + H_2 \rightarrow PH_5$, $SH_2 + H_2 \rightarrow SH_4$, $SH_4 + H_2 \rightarrow SH_6$, $ClH + H_2 \rightarrow ClH_3$, and $ClH_3 + H_2 \rightarrow ClH_5$ as well as the analogous reactions for adding two hydrogen atoms. In all cases the products are of higher energy than reactants except for the addition of atomic hydrogen to PH_3, SH_2, and SH_4. The energy difference was largest for addition of H_2 to the chlorine compounds, with ΔH^o_{298} being 477 and 503 kJ/mol for the formation of ClH_3 and ClH_5 respectively as computed in the larger basis set at the MP4 level. Thus the chlorine hydrides are predicted to be structurally stable but of remarkably high energy.

In this paper we introduced the multicenter partitioning theory of *ab initio* molecular energies in the SCF approximation as described above in Section III.B. We outlined the basic idea of this approach, compared it with previous approaches, and gave the explicit formulas for the one- and two-body terms. The total one- through four-center energies for each of the nine species in this study was reported, and the individual two- and three-center contributions were listed for all species except H_2. A similar analysis was performed on the (unstable) ClH_7 molecule at its most stable symmetric geometry, D_{5h}. These results were not unambiguous but nonetheless help quantify for the first time the origins of

23

structural instability. The results showed clearly that crowding *i.e.* a through-space steric interaction between substituents, is not a major contributor. Also comparisons of the results employing differing basis sets, by our research group and others, have shown that the most common formulation, based on the role of low-lying unoccupied d orbitals, is also not tenable. Our results showed rather that the instability of ClH_7 is due primarily to the small value of the H–Cl–H angles, analogous to ring-strain in cyclic compounds. However a more surprising finding was that the individual Cl–H two-center stabilization energies *increase* upon going from ClH to ClH_3, ClH_5, and ClH_7. This suggests that compounds analogous to ClH_7 might be exceptionally stable structurally if the H–Cl–H destabilization were lessened.

The higher chlorine hydrides are of particular interest since there had been no experimental or prior theoretical evidence for the structural stabilities of species of this type. ClH_3 exhibits a C_{2v} symmetry, while ClH_5 is C_{4v}. Their computed equilibrium structures drawn to scale, are shown below.
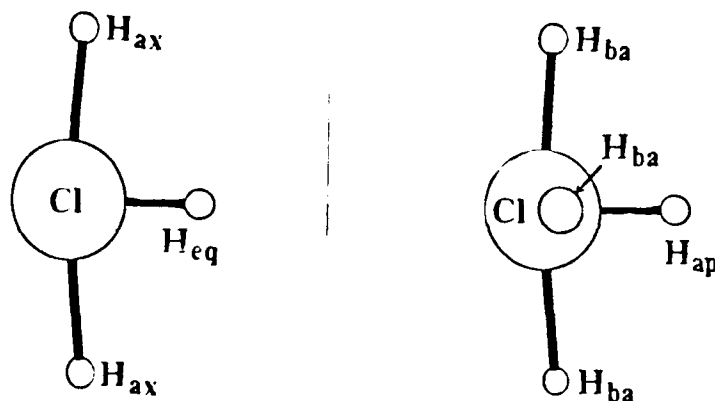
Here ClH_3 is shown on the left and ClH_5 on the right. Note that on the left one hydrogen is obscured by the chlorine.

This study demonstrated, employing these relatively simple hydrogen compounds, that *(1)* our computational procedures including MP2 structures and vibrational frequencies were feasible and useful tools and *(2)* the multicenter partitioning analysis was practicable and gave meaningful results. All these results have been published.[9]

B. NF_nH_{5-n} for $n = 3$ to 5.

Our next study in this series concerned the possible existence of highly coordinated compounds of first-row elements. Coordination to more than four substituents is well known among compounds of the second and higher rows of the periodic table. However, with the exception of some relatively weak complexes and metallic cage compounds (such as "hyperlithiated" ones[10]) no species with more than four linkages to a first-row element were known on either experimental or theoretical grounds.

24

After a large number of unsuccessful attempts which, due to the requisite computation times were concerned primarily with hydrogen substituent, we discovered a series of remarkable compounds containing fluorine. The first of these was NF_5, which exhibits a trigonal bipyramidal structure. Its structural stability was confirmed by computing all its frequencies at both the SCF and MP2 levels in the 6-31G* basis. At that time this was one of the largest (and to our knowledge the largest) computations in terms of number of basis functions to give an MP2 vibrational spectrum. Further searches revealed two more stable species, NF_4H and NF_3H_2. In both these cases the hydrogens are in the equatorial positions, indicating that although they correspond to energy minima they could decompose via a pseudorotation.

The computed equilibrium molecular structures, at the MP2 level in the 6-31G** basis, for NF_4H and NF_3H_2 are shown below. As a check the MP2 frequencies of the latter were recomputed in the 6-31++G** basis, and found to be nearly the same. Both species exhibit nearly the trigonal bipyramidal coordination of NF_5, which is of precisely D_{3h} symmetry.
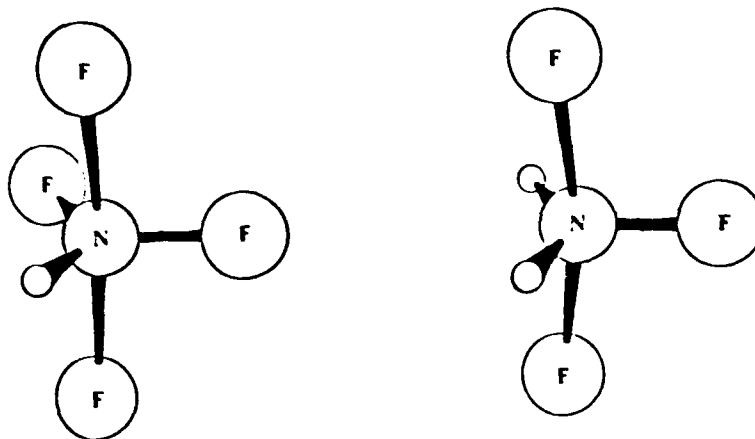
Total energies, equilibrium molecular structures, and nuclear-motion corrections to the enthalpy and free energy, ΔH°_{298} and ΔG°_{298}, were computed for these species and also, for comparison, for F_2, NF_3^{2+}, NF_3, and NF_4^+. (This was apparently the first ab initio study of NF_3^+ and only the second for NF_4^+.) Comparison of the computed structure of NF_5 with that of (planar D_{3h}) NF_3^{2+} showed that the former has distinctly covalent-type bonding and has no similarity with a more weakly bound $NF_3^{2+}/2F^-$ complex.

From the computed enthalpies and free energies, including nuclear-motion correction, we computed the thermodynamic stabilities of NF_5 relative to four possible formation or decomposition reactions. The results are summarized below in Table VI.1. These all employed the 6-31++G** basis and the MP2/6-31++G** equilibrium molecular structures. Frequencies were computed at the MP2/6-31G** level. Clearly NF_5 is stable with respect to loss of fluoride, but unstable with respect to loss of F_2. Also, the favorable

enthalpy of the third reaction is likely to complicate the synthetic preparation of NF_5 since it affords a thermodynamically favorable side reaction.

Table VI.1. Computed Reaction Enthalpies and Free Energies.

Reaction	RHF		MP2	
	ΔH°_{298}	ΔG°_{298}	ΔH°_{298}	ΔG°_{298}
$NF_4^+ + F^- \rightarrow NF_5$	−107.55	−99.81	−153.67	−145.93
$NF_3 + F_2 \rightarrow NF_5$	73.28	84.96	30.06	41.75
$NF_4^+ + F^- \rightarrow NF_3 + F_2$	−180.83	−184.77	−183.74	−187.68
$NF_3^{2+} + 2F^- \rightarrow NF_5$	−602.72	−585.46	−622.28	−605.02

The multicenter analysis was applied to the above species and, as a comparison, with the (unstable) NH_5 molecule. The origins of the stability of the fluorine compounds and the instability of NH_5 are seen to be more complex in this case than for the second-row hydrides, but it is clear that the stability of NF_5 is due to the strength of the N–F interactions, particularly the bonding to the axial flourines. This stabilization is lacking for the axial hydrogens in NH_5. Although both the "through-space" F–F repulsion in NF_5 is greater than the H–H repulsion in NH_5, and the "through-bond" is also markedly greater in the former, the N–F bonding interactions strong enough to easily overcome both these effects.

As was seen in the case of the chlorine hydrides, going from lower to higher coordination leads to stronger two-center N–F bonding interaction in the series NF_3H_2 to NF_4H to NF_5. This suggested that nitrogen compounds with coordination of degree even higher than five might also be structurally stable. Therefore we examined the possible stability of the product of fluoride addition to NF_5, NF_6^-, as discussed in the next section.

A detailed account of our work on the pentaccordianted first-row elements has recently been published.[11]

C. OF_6 and the NF_6^- and CF_6^{2-} Anions

After several trials involving nitrogen fluorides and isoelectronic fluorides of other elements, we discovered an extremely unusual molecular structure with formula NF_6^-, or hexafluoronitrate. Its structure, energy, and vibrational frequencies were computed at the SCF/6-31G*, MP2/6-31G* and MP2/6-31+G* levels. All vibrational frequencies are real in all three approximations. It was found to be precisely octahedral, with six fairly short N–F bonds (1.5848 Å at the MP2/6-31+G* level). Several related species were examined

carefully and found to be unstable. These included BF_6^{3-} (electronically unstable), as well as NeF_6^{2+}, F_7^+, NF_5H^- and CF_5H^{2-}, which are all structurally unstable. However we determined that two other octahedral hexafluorides, OF_6 and CF_6^{2-}, are electronically and structurally stable at these same levels of approximation. Detailed examination of the wavefunction for OF_6 showed that it was strongly spin-contaminated (with a much lower energy in the unrestricted than restricted Hartree-Fock approximation even though it has an even number of electrons.) OF_6 is therefore best treated by a higher-level method such as multiconfiguration SCF (MCSCF) and we have not examined its properties further. It will be the subject of a later study. However it is interesting to note that if OF_6 could be prepared, it would be an oxidant of remarkable power, giving off four fluorine atoms per molecule plus oxygen difluoride, which is highly reactive in itself.

For comparison we computed the structures, energies and vibrational frequencies of F_2, F_3^-, CF_3^-, NF_3, CF_4, NF_4^+ at the same levels of approximation. These are all experimentally characterized species except F_3^-, whose existence is still the subject of controversy. The N–F and C–F internuclear distances in NF_6^- and CF_6^{2-} were found to be comparable although slightly longer than in the smaller species. Also from the computed energies and frequencies we computed the enthalpies and free energies at 298K for several reactions related to the possible formation of NF_6^- and CF_6^{2-}. These are listed below as Table VI.2.

Table VI.2. Computed Enthalpies and Free Energies[a] at 298.15 K for Reactions Relevant to the Formation of CF_6^{2-} and NF_6^- (Kcal/mol).

		RHF/6-31G*	MP2/6-31G*	MP2/6-31+G*
(i)	$F^- + F_2 \to F_3^-$	— (—)	–46.1 (–38.4)	–27.5 (–19.7)
(ii)	$NF_4^+ + F^- \to NF_5$	–147.4 (–139.7)	–194.9 (–187.1)	–153.7 (–145.9)
(iii)	$NF_5 + F^- \to NF_6^-$	–23.9 (–13.8)	–68.9 (–58.8)	–40.0 (–29.9)
(iv)	$NF_3 + F_2 \to NF_5$	72.9 (84.5)	30.6 (42.3)	30.1 (41.7)
(v)	$NF_3 + F_2 + F^- \to NF_6^-$	49.0 (70.8)	–38.3 (–16.5)	–9.9 (11.9)
(vi)	$CF_4 + 2F^- \to CF_6^{2-}$	126.8 (144.4)	81.7 (99.4)	127.4 (145.1)
(vii)	$CF_3^- + F_2 + F^- \to CF_6^{2-}$	4.1 (25.9)	–29.0 (–7.2)	–6.4 (15.4)

These were computed at the commensurate equilibrium molecular structures and including nuclear-motion corrections computed in the MP2/6-31G* approximation. ΔG°_{298} values are given in parentheses. The results of reaction ii are the same as in section VI.B. The computed thermodynamics of the higher nitrogen fluorides, NF_5 and NF_6^-, is summarized in the following diagram. Note that, with the exception of the postulated F_3^-, NF_6^- is the

27

most thermodynamically stable in terms of its enthalpy at 298K of all the nitrogen fluorides. But as seen in Table VI.2., CF_6^{2-} is quite high in energy and is thermodynamically unstable.
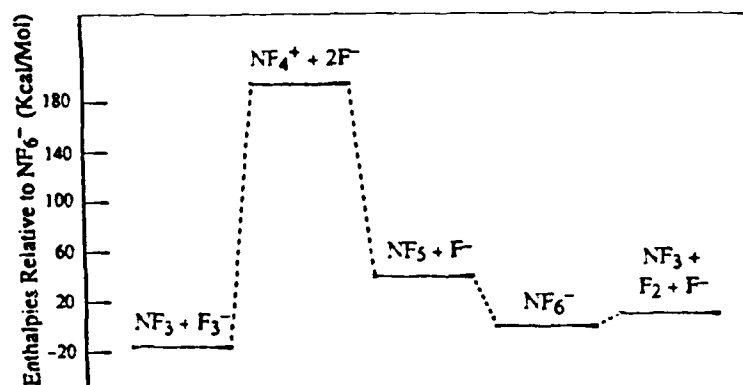
Multicenter analysis of the SCF energies for each computed species showed that, as expected, the N–F linkages in NF_6^- are very stable although the C–F two-center bonding in CF_6^{2-} is stronger yet. It is similar among all the nitrogen-flourine species and among the carbon-fluorine species. Not surprisingly, given the high degree of coordination, both NF_6^- and CF_6^{2-} have relatively large two-center "through-space" and three-center "through–bond" repulsions that destabilize their structures. However it is offset by the strength of the bonds to fluorine.

Detailed analyses of the electron distributions in all species in terms of both the atomic populations and electron topologies showed that both NF_6^- and CF_6^{2-} have strong covalent interactions between the substituents and the central atom which are very similar to those in the lower fluorides.

This was the first evidence, either experimental or theoretical, of six strong and distinct bonds to a first-row element. CF_6^{2-} is the first example of more than four such bonds to carbon. It seems likely that even more novel types of molecular structures are possible. In summary we believe that as of this writing quantum-theoretical techniques are just beginning to form a true computational chemistry in which the existence and possible means of preparing new compounds will be possible without any *a priori* experimental input. Ultimately, this must greatly accelerate the discovery of new compounds and materials with energies and other properties of immediate practical application.

VII. Stabilization of Electronic States in Crystalline Solids

A. Identification of Lattice-Stabilized CN⁻ Electronic Spectra

A segment of the overall research project quite separate from the work described above concerns the first *ab initio* computation of stabilization by crystalline solids of excited

electronic states of anions, states that are electronically unstable with respect to electron autodetachment *in vacuo*. It originated in part from the ideas set forth in section II.C of the original research proposal.

It has been known experimentally for some time that the surfaces of many types of alkali halide crystals emit a prominent "molecular" spectrum in the 2200–3600 Å region following excitation by electrons, ions or vacuum ultraviolet radiation. Although this emission was found even from the purest available crystals and its origin was long considered a mystery, doping and similar experiments pointed to some electronically excited state of CN^- (apparently often present as an impurity) as the emitting species. Recent studies using synchrotron radiation, along with isotopic enrichment, concluded the CN^- was involved and gave a sufficiently detailed spectrum that accurate spectroscopic constants could be determined. In addition it was shown that, by releasing the energy of the exciting beam as photons, the CN^- impurity centers effectively shielded the crystals from radiation damage, suggesting ultimate practical applications.

From the point of view of molecular quantum mechanics, this is a particularly intriguing phenomenon since small anions cannot have excited electronic states *in vacuo* since the energy of excitation is considerably greater than the energy to strip off an electron. In other (more theoretical) words the wavefunctions of all excited states are arbitrary with respect to any admixture of continuum wavefunctions. However there remained the possibility that in ionic crystals such as the alkali halides, the Madelung potential of the array of localized charges formed by the ions of the lattice could stabilize the energy of the anion relative to the neutral CN by more than the electron-excitation energy. This in turn would allow a complete spectrum of transitions to excited states where none existed for the free anion.

Since precise data for CN^- had become available, we examined the technical questions of *(1)* what theoretical technique would be practicable for computing the spectroscopic parameters (internuclear distance, vibrational frequencies, anharmonicity, *etc.*) with sufficient reliability and accuracy, and *(2)* what if any computational models specifically for the infinite alkali-halide lattice would be both practicable and give realistic results. We would then be able to address the assignment of the observed emission spectrum.

We determined that an MCSCF (multiconfiguration self-consistent field) approach employing the 6-311++G(2d) basis was sufficiently accurate and flexible for application to the range of states and environments we needed to study. We verified this by test computations on CN, whose spectrum in the gas phase is well characterized and could by used for comparison. Also by repeated test calculations we showed that in particular this procedure is capable of giving extremely accurate relative energies between electronic states. We then examined a series of computational models for the ionic lattice. Since three shells of cations, anions, and cations define the three symmetric orientations (100, 110 and 111 directions) in a face-centered cubic lattice such as characterize most alkali halides, and

since this also gives approximate charge neutrality (net charge of +1) we used this array of 26 point charges to represent the ionic lattice.

The experimental data could be interpreted by in terms of the excitation energy to the observed excited state and the associated increase in the equilibrium bond length. We computed these for three electronic states that are the most stable in isoelectronic species: $^3\Sigma^+$, $^1\Pi$, and $^3\Pi$. From initial computations varying the internuclear distance of each electronic state along the (100) direction it was immediately clear that our computed spectroscopic constants for the $^3\Sigma^+$ state agreed well with experiment. The Π states however had excitation energies too high and internuclear distances far too small to account for the observed spectra. Upon going to the (110) orientation, which more closely approximates the experimental average equilibrium orientation, the agreement with the computed $^3\Sigma^+$ state became even better. In this case computed the bond length in the ground state to be 1.184 Å, in the $^3\Sigma^+$ state 1.358 Å and the excitation energy to be 45300 cm^{-1}. These agree well with the observed bond length increase of +0.17 Å and excitation energy of 45350 cm^{-1}. On the other hand the $^3\Pi$ state gives an internuclear distance of only 1.280 Å and excitation energy of 57400 cm^{-1}. The difference between the results we obtain for the (100) and (110) directions in the simulated lattice imply that there is a marked orientation preference by the excited state.

From this study we were able to demonstrate that the emission from excited alkali-halide surfaces is definitely due to CN$^-$ and the emitting state is $^3\Sigma^+$. Perhaps more importantly this work showed for the first time that *ab initio* computations based on a fairly modest representation of the infinite ionic lattice could give results of quite useful accuracy for predicting the properties of solids. These initial results were subsequently published.[12]

B. Precise Computation of Lattice Effects

Once it was demonstrated that *ab initio* computations are useful in understanding the properties, particularly in assigning the spectra, of the anionic impurity centers such as CN$^-$ we examined the accuracy of various models for the solid lattice in considerable detail. Specifically we studied for the ground state *(1)* the forms of SCF versus MCSCF energy curves, *(2)* the effects of varying the lattice constant from that in LiCl to NaCl, KCl, and NaBr, *(3)* the effects of changing the number of point charges in the lattice from 26 to as few as 6 (only the first shell of charge), *(4)* the improvement in accuracy upon replacing the point charges by "real ions" computed in the SCF approximation with a minimal basis set for the first shell of ions in LiCl, NaCl, KCl, and NaBr, and *(5)* whether this model is substantially improved by increasing the size of the basis to 6-31G on the "real ions" in the case of the LiCl lattice. We also studied carefully the ways in which the computed discrete MCSCF energies are most accurately and reliably analyzed in terms of the spectroscopic parameters.

It was found that the spectroscopic properties of the ground state were remarkably insensitive to the precise nature of the model used to emulate the infinite lattice. For

example there was relatively little difference on going from six to 26 point charges. Employing the "real ions" in each case caused a compression of the bond length with concomitant increase in vibrational frequency by almost exactly the correct amounts (typically on the order of 0.01 Å and 100 cm^{-1} respectively) to bring our MCSCF results into agreement with the experimental data for the ground state. However the use of basis sets larger than the minimal set caused much smaller changes, with differences in CN$^-$ bond length and vibrational frequency of only 0.001 Å and 10 cm^{-1}.

These results all augur well for the rapid development of *ab initio* computations of impurity centers and other defects in solids, since we have now demonstrated conclusively that at least for the spectroscopic parameters of CN$^-$ in ionic lattices as a prototype fairly simple computational models can be made to yield results of remarkable accuracy and consistency. However related properties we have examined less carefully, such as the location and preferred orientations of the diatomic anion inside the defect, give less satisfactory results at this level of approximation and will be the subject of several later studies. In summary, we believe that we have only opened the book, rather than closed it, for continued research in this several related areas.

Finally, we not that this more comprehensive study has recently been published.[13]

VIII. References Cited

1. The best summary of this work covering approximately a six-year period is given by K. McDowell and L. Lewis, *J. Chem. Phys.* **77**, 3068 (1982).
2. C. S. Ewig, L. Smentek-Mielczarek and B. A. Hess, Jr., *Theoret. Chim. Acta,* **75**, 129 (1989).
3. E. Clementi and W. von Niessen, *J. Chem. Phys.*, **54**, 521 (1971).
4. H. Tatewaki and S. Huzinaga, *J. Comput. Chem.* , **1**, 205 (1980).
5. R. Ditchfield, *Mol. Phys.* , **27**, 789 (1974). See also C. McMichael Rohlfing, L. C. Allen, and R. Ditchfield, *J. Chem. Phys.*, **63**, 185 (1981).
6. U. Fleischer, M. Schindler, and W. Kutzelnigg, *J. Chem. Phys.* , **86**, 6337 (1987).
7. J. R. Van Wazer, C. S. Ewig, and R. Ditchfield, *J. Phys. Chem.* **93**, 2222 (1989).
8. G. del Conde, P. S. Bagus, and C. W. Bauschlicher, Jr., *Theor. Chim. Acta* **45**, 121 (1977).
9. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.*, **111**, 1552 (1989).
10. P. v. R. Schleyer, E. U. Würthwein, E. Kaufman, T. Clark, and J. A. Pople, *J. Am. Chem. Soc.*, **105**, 5930 (1983).
11. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.*, **111**, 4172 (1989).
12. C. S. Ewig and J. Tellinghuisen, *Chem. Phys. Letters*, **153**, 160 (1988).
13. J. Tellinghuisen and C. S. Ewig, *J. Chem. Phys.* **91**, 5476 (1989).

IX. Publications Resulting from This Grant

1. P. L. Polavarapu, C. S. Ewig and T. Chandramouly, *J. Am Chem. Soc.*, **109**, 7382 (1987). Conformations of Tartaric Acid and its Esters.

2. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.*, **110**, 79 (1988). *Ab Initio* Structures of Phosphorus Acids and Esters. 3. The P-O-P Bridged Compounds $H_4P_2O_{2n-1}$ for n=1 to 4.

3. D. Michalska, L. J. Schaad, P. Cársky, B. Andes Hess, Jr. and C. S. Ewig, *J. Comput. Chem.*, **9**, 495 (1988). Basis Set Effects and the Choice of a Reference Geometry in *Ab Initio* Calculations of Vibrational Spectra.

5. C. S. Ewig and J. Tellinghuisen, *Chem. Phys. Letters*, **153**, 160 (1988). Ab Initio Study of Excited States of CN$^-$ Stabilized in Point-Charge Lattices.

6. C. S. Ewig, L. Smentek-Mielczarek and B. A. Hess, Jr., *Theoret. Chim. Acta,* **75**, 129 (1989). Perturbative Corrections to Basis Incompleteness In Molecular SCF Calculations.

7. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.*, **111**, 1552 (1989). Ab Initio Studies of Molecular Structures and Energetics. 2. Energies and Stabilities of PH_n, SH_n and ClH_n Compounds.

8. J. R. Van Wazer, Carl S. Ewig and Robert Ditchfield, *J. Phys. Chem.,* **93**, 2222 (1989). *Ab Initio* Computation of Silicon-29 Nuclear-Magnetic Resonance Chemical Shifts.

9. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.*, **111**, 4172 (1989). Ab Initio Studies of Molecular Structures and Energetics. 3. Pentacoordinated NF_nH_{5-n} Compounds.

10. P. K. Bose, D. O. Henderson, C. S. Ewig and P. L. Polavarapu, *J. Phys. Chem.*, **93**, 5070 (1989). Experimental and Ab Initio Spectra of 1,2–Dibromoethane, Meso–1,2–Dideutero–1,2–Dibromoethane and Chiral–1,2–Dideutero–1,2–Dibromoethane.

11. J. Tellinghuisen and C. S. Ewig, *J. Chem. Phys.* **91**, 5476 (1989). Ab Initio Study of CN$^-$ Impurity Centers in Alkali Halides: Lattice Stabilization of Excited Electronic States.

12. C. S. Ewig and J. R. Van Wazer, *J. Am. Chem. Soc.* xxx, xxxx (1989). Ab Initio Studies of Molecular Structures and Energetics. 4. The Hexacoordinated CF_6^{2-} and NF_6^- Anions.

13. J. R. Van Wazer, V. Kellö, B. A. Hess, Jr., and C. S. Ewig (*J. Phys. Chem.*, submitted for publication). Accuracy in Ab Initio Reaction-Energy Computations. I. Compounds of First-Row Elements.

14. C. S. Ewig (*J. Chem. Phys.*, submitted for publication). The Theory of Multicenter Partitioning of Molecular Energies.

15. C. S. Ewig, B. A. Hess, Jr., V. Kellö, and J. R. Van Wazer (in preparation) The Accuracy of Coupled–Cluster Theory in the Computation of Molecular Enthalpies.

The Theory of Multicenter Partitioning of Molecular Energies

Carl S. Ewig

Department of Chemistry
Vanderbilt University
Nashville, Tennessee 37235

Abstract

The theory of analyzing molecular energies, as found from *ab initio* computations, in terms of the contributions from specific sets of atoms or centers is discussed. It is shown that the basis-function expansion of molecular wavefunctions in general leads to energetic interactions involving only one- through four-center terms. For the special case of SCF energies the formulas for the energy terms are given explicitly and several numerical properties are presented, including basis-set dependence, the differences between terms that arise in chemical reactions, and the first and second derivatives of individual terms with respect to molecular deformations. Expressions are also given for the multicenter resolution of correlation energies employing second-order many-body-perturbation theory. Some numerical examples of the latter are also presented.

I. Introduction

In the interpretation of the general features of molecular structures and energetics, as studied on either an experimental or theoretical level, there is often recourse to descriptions invoking the interactions between specific individual atoms or sets of atoms. Many commonly accepted rationales invoke the presence of strong bonding interactions between specific atoms, interactions through space between atoms not considered to be bonded, as well as multicenter bonding, high-energy bonds, *etc.* However to our knowledge no general theoretical approach has been developed for testing a given molecular electronic structure to determine the pattern and extent with which the energy is distributed among atoms.

This paper discusses the partitioning of molecular energies in cases where the wavefunction is described by a linear combination of atom-centered basis functions. The energy corresponding to any such wavefunction may be written in terms of integrals centered on one nucleus only, specific pairs of nuclei, *etc.* The premise is that the energy associated with one of these sets of centers is that described by the basis functions having their origins on these centers. This is to some extent arbitrary, since each basis function assumes nonzero values throughout the molecule (except on nodal surfaces). However, as has been pointed out[1], it is no more or less arbitrary than discussing molecules in terms of their constituent atoms.

In a previous paper[2] we discussed the use of the multicenter resolution of the SCF (self-consistent-field or Hartree-Fock) energy to examine multicenter bonding in highly coordinated compounds containing third-period elements, and presented explicit expressions for the one- and two-center terms, as well as the resulting numerical values. We have also reported computed SCF results of this type for lithium, carbon and nitrogen compounds.[3,4] In this paper we examine the theoretical framework for the general multicenter resolution of molecular energies, examine the SCF case in particular detail, and give both analytic expressions for the individual terms as well as numerical examples for several properties of compounds containing second-period elements. Finally the analogous expressions for including electron correlation corrections at the second-order many-body-perturbation level are presented, along with representative numerical results for H_2 and H_2O.

II. General Properties

The basic principle is that the energy of any system, with a specific geometric and electronic structure, may be expanded in terms of individual energies associated with centers or sets of centers (generally atomic nuclei),

$$E^{Total} = \sum_n E_n^{(1)} + \sum_{mn} E_{mn}^{(2)} + \sum_{lmn} E_{lmn}^{(3)} + ... \tag{1}$$

where the summations are over all centers in the system. The $E_n^{(1)}$ will be referred to as one-center energies, the $E_{mn}^{(2)}$ as two-center energies, *etc.*

It is easily seen that the number of such terms increases rapidly as the number of centers becomes large.[5] If there are N atomic centers, the number of m-center terms, N_m is simply

$$N_m = \frac{N!}{m! \, (N-m)!}$$

and so as $N \rightarrow \infty$, $N_m \rightarrow N^m/m!$. Since the number of E_{m+1} terms will then be N/m times the number of E_m terms, this suggests that as the size of a molecular system increases and $N \rightarrow \infty$ (as in infinite solids), the sum of all the E_{m+1} must dominate the sum of the E_m, the E_{m-1}, *etc.* (This might be termed the "multicenter catastrophe.") However, as shown below, in practice the average value of each of the E_m is found to decrease very rapidly with increasing m, and the expansion (1) converges rapidly.

In fact it may be shown that the energy expansion terms are all exactly zero for all m greater than four. This may be readily seen from the expression for the energy based on the electronic operator employing the second-quantization approach,[6,7]

$$H = \sum_r \sum_s h_{rs} a_r^\dagger a_s + \frac{1}{2} \sum_r \sum_s \sum_t \sum_u g_{rstu} a_r^\dagger a_t^\dagger a_u a_s$$

where h_{rs} is the one-electron Hamiltonian consisting of kinetic-energy and electron-nuclear-attraction terms, g_{rstu} is a two-electron repulsion integral, the functions r, s, t, and u are molecular orbitals, and the $a^\dagger$ and a are the spin-orbital creation and annihilation operators respectively. Without significant loss of generality we can expand the molecular wavefunction in terms of determinantal functions with normalized expansion coefficients c_n. Evaluating the Hamiltonian over this wavefunction to obtain the energy E,

$$E = \sum_r \sum_s h_{rs} \gamma_{rs} + \frac{1}{2} \sum_r \sum_s \sum_t \sum_u g_{rstu} \Gamma_{rstu} \qquad (2)$$

where γ_{rs} and Γ_{rstu} are the one- and two-particle density matrices respectively,

$$\gamma_{rs} \equiv \sum_n \sum_{n'} c^*_n \langle n | a_r^\dagger a_s | n' \rangle c_{n'},$$

$$\Gamma_{rstu} \equiv \sum_n \sum_{n'} c^*_n \langle n | a_r^\dagger a_t^\dagger a_u a_s | n' \rangle c_{n'}$$

Note that for a given set of determinantal functions n with coefficients c_n the density matrix elements do not contain any of the physical parameters of the system, such as interatomic distances or nuclear charges. All such descriptors upon which the energy depends are of course contained in the h_{rs} and g_{rstu}. As a physical analogy, if it were possible to place an imaginary, perfect, dielectric barrier separating two atoms in such a way that the wavefunction remained unchanged, then γ_{rs} and Γ_{rstu} would not be affected and hence they would not alter the total energy. (The same principle clearly applies to the more general case of imaginary partitions between atoms that alter the wavefunction but only in such a way as to leave the coefficients c_n and hence the density matrices unchanged.) Therefore we will consider only the h_{rs} and g_{rstu}.

If each of the functions is expanded in a set of analytic basis functions $| n \rangle = \sum_i C_{ni} | i \rangle$ then h_{rs} will contain terms involving integration over pairs of basis functions, $h_{rs} = \sum_i \sum_j C_{ri} C_{sj} h_{ij}$. Similarly g_{rstu} contains integrals over sets of four basis functions at a time, all of which may be the same or differing in any combination. We may now

segregate all such terms in (2) that contain integrals over only functions centered at one point (typically a nucleus) m. Its contribution to the energy is $E_m^{(1)}$. Thus, for example,

$$E_n^{(1)} = \sum_{\substack{ij \text{ on } n}} \sum_r \sum_s C_{ri} C_{sj} h_{ij} \gamma_{rs} +$$

$$+ \sum_{\substack{ijkl \text{ on } n}} \sum_r \sum_s \sum_t \sum_u C_{ri} C_{sj} C_{tk} C_{ul} g_{ijkl} \Gamma_{rstu}.$$

Clearly h_{rs} may contribute to $E_m^{(1)}$ and $E_{mn}^{(2)}$ while g_{rstu} may contribute one- through four-center terms. However it is not possible to obtain five- or higher-center terms.

The complete Hamiltonian is obtained by adding the nuclear-nuclear repulsion terms $Z_m Z_n / R_{mn}$ where Z_m is the charge on nucleus m and R_{mn} is the internuclear distance. These must be added to the corresponding $E^{(2)}$ terms to obtain two-center part of the total energy.

III. The Hartree-Fock Case

In the special case where the molecular wavefunction contains only a single determinantal function with doubly occupied orbitals, both γ_{rs} and Γ_{rstu} in equation (2) may be greatly simplified, since

$$\gamma_{rs} = 2\delta_{rs}$$

$$\Gamma_{rstu} = 4\delta_{rs}\delta_{tu} - \delta_{rt}\delta_{su} - \delta_{ru}\delta_{st}$$

Substituting into (2) and writing the orbital expansion coefficients in terms of the density matrix over basis functions $d_{ij} \equiv 2 \sum_n C_{ni} C_{nj}$ we obtain the analytic expressions for each term of the multicenter expansion. $E^{(1)}$ contains kinetic, electron-nuclear attraction and electron-electron repulsion terms,

$$E_n^{(1)} = \sum_{\substack{ij \text{ on } n}} d_{ij} (<i|-\nabla^2/2|j> + <i|-Z_n/r|j>) + 1/4 \sum_{\substack{ijkl \text{ on } n}} g_{ijkl} (2d_{ij}d_{kl} - d_{ik}d_{jl}).$$

$E_{mn}^{(2)}$ consists of kinetic, electron-nuclear attraction, electron-electron repulsion and the nuclear-nuclear repulsion terms,

5

$$E_{mn}^{(2)} = 2 \sum_{i \text{ on } n} \sum_{j \text{ on } m \neq n} d_{ij} (<i|-\nabla^2/2|j> + <i|-Z_m/r|j> + <i|-Z_n/r|j>) +$$

$$\sum_{ij \text{ on } m \neq n} d_{ij} <i|-Z_n/r|j> + \sum_{ij \text{ on } n \neq m} d_{ij} <i|-Z_m/r|j> +$$

$$1/4 \sum_{ijkl \text{ on } n \, m \neq n} g_{ijkl} (2d_{ij}d_{kl} - d_{ik}d_{jl}) + Z_m Z_n/R_{mn}$$

Similarly $E_{lmn}^{(3)}$ = electron-nuclear attraction and the two-electron terms, and $E_{klmn}^{(4)}$ contains

two-electron terms only.

That the SCF energy could be decomposed in this fashion was first pointed out two decades ago by Clementi and coworkers.[8-12] As an example of the relative magnitudes involved, we list in Table I each of these terms that arise for H_2O. These were computed in the 6-311G(2d,p) basis[13] which includes two d functions on the first-period atom and one p function on each hydrogen. The bond lengths and angles of this and all subsequent species were taken from the experimental equilibrium molecular structures.[14] The contributions to the one-center terms are seen to approximate those of the free atoms. Although the kinetic energy and the Coulomb integrals in the two-electron energy are positive and hence destabilizing, the negative electron-nuclear attraction dominates. The kinetic energy is a relatively small contributor to the two-center energy, which is essentially a compromise between the stabilizing electron-nuclear attraction and the destabilizing electron-electron and nuclear-nuclear repulsion. This is of course in accord with the conventional picture of a "covalent bond." The "bonded" H–O interactions are negative (–0.84714 hartrees) and hence stabilizing with respect to the total energy, while the H–H terms are seen to be small and positive and hence slightly repulsive. The H–O–H term, –0.03916 hartrees, represents a very modest three-center bonding energy. The total one-center energy, –74.47128 is larger by approximately a factor of 50 than the total two-center contribution, which in turn is about a factor of 50 larger than $E^{(3)}$.

Table II shows the extent to which the individual $E^{(1)}$ and $E^{(2)}$ terms depend on the choice of basis set for four representative diatomic compounds: H_2, Li_2, N_2 and F_2 in each of four basis sets: 3-21G,[15] 6-31G,[16] 6-31G(d)[17] and 6-311G(2d).[13] From this table it is clear that the basis-set dependence is slight, differences in $E^{(1)}$ and $E^{(2)}$ among the various basis sets being approximately the same magnitude as the differences in their total energies.

6

The total $E^{(1)}$, $E^{(2)}$ and $E^{(3)}$ for H_2O shown in Table I are compared in Table III with the analogous terms for four additional polyatomic species – N_2O, CO_2, NH_3 and NF_3 – thus demonstrating the convergence of the total energy with number of terms in the multicenter expansion, (1). As usual, $E^{(2)}$ is negative and hence stabilizing. $E^{(3)}$ is positive and hence destabilizing in each case shown except H_2O. However it is only significant for the tabulated species not containing hydrogen. Both NH_3 and NF_3 exhibit a slight four-center stabilization.

A. Reaction Energies

From Eq. (1) and the derivation above it is clear that if a molecule is formed from its constituent atoms, for which of course only $E^{(1)}$ is nonzero, the total energy of the reaction will be

$$\Delta E^{Total} = \sum_n (E_n^{(1)} - \varepsilon_n) + \sum_{mn} E_{mn}^{(2)} + \sum_{lmn} E_{lmn}^{(3)} + ...$$

where ε_n is the energy of the isolated atom n. This is simply the energy of deforming the individual atoms (the first term), plus the pair-wise stabilization or destabilization (second term), *etc.*[12] For example Table IV lists these terms for the reaction $2H + O \rightarrow H_2O$. This is the reaction employed by Pauling[18] to derive the O–H "bond energy" which he took to be the total enthalpy of this reaction divided by two. Such "bond energies" are often transferrable among similar molecules, and are still widely employed. From Table IV it may be seen that O–H two-center energies are indeed the largest and negative, and hence stabilizing. However they are much larger than the value postulated by Pauling based on the experimental reaction enthalpy of –221.14 kcal/mole. (The currently accepted value of this enthalpy difference is -219.352 ± 0.026.[19]) Note that there are also there are large positive contributions from $\Delta E^{(1)}$. This is of course in accord with the finding of modern electronic-structure theory that *total* molecular energies are much more sensitive to changes in the wavefunctions near the nuclei than in the bonding regions between the nuclei.

Table V shows a similar comparison for reactions between diatomic molecules. The differences in the total two-center energies $\Delta E^{(2)}$ between products and reactants is generally the dominant term, although differences in one-center energies are always large and may, as in the case of $H_2 + Li_2 \rightarrow 2LiH$, be as large as $\Delta E^{(2)}$.

7

B. Derivatives With Respect to Molecular Deformations

A second property of the multicenter expansion that follows from our derivation above is that if λ is any parameter (such as nuclear charge, nuclear coordinates, or electronic charge density) that gauges the effect of an individual atom or set of atoms on the overall molecular energy, then

$$\frac{\partial E^{(m)}}{\partial \lambda}\bigg|_C = \frac{\partial E^{Total}}{\partial \lambda}\bigg|_C$$

for some value of m between one and four. Here in taking the derivatives the orbital expansion coefficients C (and hence the density matrices) are to be held constant. For example in a diatomic molecule,

$$\frac{\partial E^{(2)}}{\partial R}\bigg|_C = \frac{\partial E^{Total}}{\partial R}\bigg|_C$$

where R is the internuclear distance. Similar relations hold for the second through fourth derivatives, the higher derivatives being zero. Of course when the wavefunction is also allowed to vary, these are numerically quite different.

In Table VI we list the first and second derivatives of $E^{(1)}$ and $E^{(2)}$ as well as E^{Total} for ten representative diatomic species at their experimental equilibrium geometries. These were evaluated by computing each term in the energy expansion at internuclear distances displaced from the experimental equilibrium by ± 0.01 Å. In evaluating $\partial E^{Total}/\partial R$ (= $\partial E^{(1)}/\partial R + \partial E^{(2)}/\partial R$) and $\partial^2 E^{Total}/\partial R^2$ (= $\partial^2 E^{(1)}/\partial R + \partial^2 E^{(2)}/\partial R$) the wavefunction was allowed to vary, so the differences between the second and third columns of data reflect the effect of relaxation of the wavefunction with changes in R. It is clear that contributions of the one-center energies are again of nearly the same magnitudes as those of the two-center energies. However, as we have generally found to be the case, $\partial^2 E^{(2)}/\partial R$ tends to be the dominant contributor to $\partial^2 E^{Total}/\partial R^2$.

The same analysis may be applied to polyatomic species. In Table VII are listed the derivatives of $E^{(1)}$, $E^{(2)}$ and $E^{(3)}$ of H_2O with respect to α, the HOH bond angle, and R, the symmetric OH bond stretching symmetry coordinate. $R = 2^{-1/2}(\delta R_{OH1} + \delta R_{OH2})$ where R_{OH1} and R_{OH2} are the O–H internuclear distances. The most important term in both derivatives is the one-center oxygen energy, which is largely negated by an OH two-center contribution of opposite sign. The role of the H–H two-center energy and of the three-center energy is relatively small, although large compared with the total derivatives. We conclude that, as in the case of reaction energies, the derivatives tend to exhibit very

8

significant contributions of terms, particularly one-center energies, which are different from those that might be expected from conventional ideas of chemical bonding.

IV. Second-Order Correlation Energies

From the discussion in Section I, the multicenter resolution of the energy may be performed for any molecular electronic description for which it is possible to express the one- and two-particle density matrices. In the case of approximation to the energy including electron-correlation through second order in the many-body perturbation theory in the Møller-Plesset sense (MP2) there will be additional terms in the multicenter expansion of the energy due to electron-electron terms in the Hamiltonian. Specifically for the case of doubly occupied orbitals we may set

$$\Gamma'_{rstu} = 4\ (g_{rstu} - g_{rsut})/(e_r + e_s - e_t - e_u)$$

where the e's are the orbital energies. The multicenter expansion then consists of terms, in addition to the SCF, of the form

$$E_n^{(1)} = \sum_{ijkl\ on\ n}\ \sum_{rstu} C_{ri}\ C_{sj}\ C_{tk}\ C_{ul}\ g_{ijkl}\ \Gamma'_{rstu}$$

$$E_{mn}^{(2)} = \sum_{ijkl\ on\ mn}\ \sum_{rstu} C_{ri}\ C_{sj}\ C_{tk}\ C_{ul}\ g_{ijkl}\ \Gamma'_{rstu}$$

and similarly for $E_{lmn}^{(3)}$ and $E_{klmn}^{(4)}$.

Not surprisingly the results of this resolution of the energy are qualitatively much different from those of the preceding section. Thus for example H_2 in the 6-311G(p) basis gives individual one-center correlation energy $E^{(1)} = -0.06071$ hartrees, while $E^{(2)} = 0.03307$ to give the total MP2 energy of -0.027798. Note that, in contrast to the SCF case, the electron-correlation $E^{(2)}$ for the H-H interaction is positive and hence destabilizing. This is consistent with the usual finding that inclusion of electron correlation lengthens optimized internuclear distances and decreases computed stretching force constants.[20]

The same type of multicenter resolution of the correlation energy is given in Table VIII for H_2O . This also employed the 6-311G(2d,p) basis. The correlation energy per hydrogen, -0.02435, is a large fraction of the analogous SCF term (-0.14617) but comparable to the value in H_2. In oxygen the correlation energy is a much smaller fraction

of the SCF energy (–74.17894). However the former is comparable to the MP2 of atomic oxygen, –0.12764.[21] Thus the MP2/SCF ratio of the oxygen $E^{(1)}$ terms in H_2O is on the same order as in the free atom. The O-H bonding interaction is again positive, while the H-H term is negative although practically zero. $E^{(3)}$, the three-center stabilization, is appreciable when compared to the analogous value in Table I.

V. Discussion

Despite the very wide-spread usage of such concepts as "high-energy bonds", "multicenter bonding", "anomeric stabilization", *etc.* there have been remarkably few attempts to quantify the *ab initio* energies of molecular systems in terms of interactions between specific atoms or centers. However one such approach, termed the "many-body expansion", has been proposed[22] and applied to several test cases.[5,22-25] In the present notation this is written

$$E^{Total} = \sum_n E_n^{(1)} + \sum_{mn} E_{mn}^{(2)}(R_{mn}) + \sum_{lmn} E_{lmn}^{(3)}(R_{lmn}) + ... \qquad (3)$$

where each energy term is associated with a characteristic functional dependence on the internuclear coordinates R_{mn}, R_{lmn}, *etc.* This clearly has several properties in common with the expansion (1) above; for example $E_{lmn}^{(3)}$ becomes zero if any of centers l, m or n is removed to an infinite separation from the others. The terms in (1) have this property not only in terms of the coordinates of the centers but also any parameter that affects the energy, such as nuclear coordinates or local charge densities. In conjunction with *ab initio* computations, however, there is a serious problem in deriving numerical values of the individual terms employing (3). Although the literature is generally not explicit, these apparently are to be obtained by dividing the molecular system into fragments and equating the energies of fragments with terms in the expansion.[5] Thus $E_n^{(1)}$ is taken to be ε_n, the energy of the lowest electronic state of the free atom n for any molecule in which it appears. $E_{mn}^{(2)}$ is assumed to be the same in all molecules containing m and n. Thus for example the N_3 molecule is represented[24] by the energies of three ^4S nitrogen atoms plus two N-N energy functions taken from the N_2 molecule. The three-body energy is then purported to be the difference between the total energy and the sum of these one- and two-body terms. However the accuracy of this type of assumption is generally difficult to verify independently. Conceptually this is of course very much along the same lines as the

10

"bond-energy" procedure of Pauling as discussed above. It thus appears to have much the same weaknesses but without the strengths.

There has been considerable conjecture for several years concerning "multicenter bonding" such as the three-center model originally proposed by Musher[26-28] which has recently been extended as far as "five-center bonding[29]." The quantitative basis for such descriptions is generally mysterious, but seems to be based on the facility with which the species exhibits delocalized *canonical* molecular orbitals. However as has been emphasized by Musher[28] in the case of three-center four-electron bonding, the delocalized canonical molecular orbitals are precisely equivalent within a localizing transformation (which leaves the energy unchanged) with two-center localized two-electron bonding orbitals.

Similarly there have been attempts to describe interaction among centers based on various qualitative descriptions of localized orbitals such as those proposed for one- and two-center analyses by Boys[30,31] and one- through three-center orbitals by Edmiston and Ruedenberg.[32,33] These have been discussed in particular by Lipscomb and coworkers.[34,35] Unfortunately these the disadvantages of considerable ambiguities in their definitions[35] and lack of a clear connection with molecular energies.[36] Nevertheless it is very interesting to note that the relative importance of one-, two- and higher-centered energies roughly parallels those of the analogous localized orbitals. Thus $E^{(3)}$ and $E^{(4)}$ are the energetic counterparts of the "localization tails" that arise from those portions of the electron distributions not accounted for by one- and two-center localized orbitals.

In this paper it has been shown that employing the two-particle density matrix the energies of molecular species may be uniquely resolved into separate terms depending on the parameters (such as positional coordinates) characterizing individual centers, pairs, triplets and quartets of centers. It applies to any computational method for which the energy expression may be generalized as in (2) above, such as general multiconfiguration SCF (MCSCF) formalism.[7] Further, we have shown that it exhibits the useful properties of rapid convergence (relatively small four-center energies with higher-center energies being identically zero) and relative independence of the choice of basis set. The present approach presupposes a particular geometry and wavefunction so that terms, such as overlap integrals or gradients, that appear in optimization of the wavefunction or geometry need not be explicitly included. By providing a quantitative picture of energy distributions within molecular species it permits a means of testing the results of quantum-theoretical computations in terms of commonly held concepts of interactions between atoms or other centers.

Acknowledgment

References

[1]E. Clementi and W. von Niessen, J. Chem. Phys. **54**, 521 (1971).

[2]C. S. Ewig and J. R. Van Wazer, J. Am. Chem. Soc. **111**, 1552 (1989).

[3]C. S. Ewig and J. R. Van Wazer, J. Am. Chem. Soc. **111**, 4172 (1989).

[4]C. S. Ewig and J. R. Van Wazer, J. Am. Chem. Soc. (in press).

[5]J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, and A. J. C. Varandas, *Molecular Potential Energy Functions* (Wiley, Chichester, 1984).

[6]P. Jørgensen and J. Simons, *Second Quantization-Based Methods in Quantum Chemistry* (Academic Press, New York, 1981).

[7]R. Shepard, in *Ab Initio Methods in Quantum Chemistry – II*, edited by K. P. Lawley (Wiley, Chichester, 1987).

[8]E. Clementi, J. Chem. Phys. **46**, 3842 (1967).

[9]E. Clementi, Intern. J. Quantum Chem. **IIIS**, 179 (1969).

[10]E. Clementi and A. Routh, Intern. J. Quantum Chem. **VI**, 525 (1972).

[11]G. Corongiu and E. Clementi, Gazz. Chim. Ital. **108**, 273 (1978).

[12]E. Clementi, *Lecture Notes in Chemistry, Vol. 19* (Springer, Berlin, 1980), Chap. 3.

[13]R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[14]M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman, D. A. Ramsay, F. J. Lovas, W. J. Lafferty, and A. G. Maki, J. Phys. Chem. Ref. Data **8**, 619 (1979).

[15]J. S. Binkley, J. A. Pople, and W. J. Hehre, J. Am. Chem. Soc. **102**, 939 (1980).

[16]W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[17]P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[18]L. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, 1960), p. 84.

[19]M. W. Chase, Jr., C. A. Davies, J. R. Downey, Jr., D. J. Frurip, R. A. McDonald, and A. N. Syverud, J. Phys. Chem. Ref. Data **14**, Suppl. 1 (1985).

[20]D. Michalska, L. J. Schaad, P. Cársky, B. A. Hess, Jr., and C. S. Ewig, J. Comput. Chem. **9**, 495 (1988).

[21]This value is based on the MP2 approximation with "unrestricted Hartree-Fock" orbitals. See J. A Pople and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1954) and R. McWeeny and G. Diercksen, J. Chem. Phys. **49**, 4852 (1968).

[22]W. Kolos, F. Nieves and O. Novaro, Chem. Phys. Letters **41**, 431 (1976).

[23]J. N. Murrell, Chem. Phys. Letters **55**, 1 (1978).

[24]J. N. Murrell, O. Novaro, S. Castillo, and V. Saunders, Chem. Phys. Letters **90**, 421 (1982).

[25]W. L. Feng, O. Novaro, and J. Garcia-Prieto, Chem. Phys. Letters **111**, 297 (1984).

[26]J. I. Musher, Angew. Chem., Int. Ed. Engl. **8**, 54 (1969).

[27]J. I. Musher, J. Am. Chem. Soc. **94**, 1370 (1972).

[28]V. B. Koutecky and J. I. Musher, Theoret. Chim. Acta (Berlin) **33**, 227 (1974).

[29]W. B. Farnham, D. A. Dixon, and J. C. Calabrese, J. Am. Chem. Soc. **110**, 8453 (1988).

[30]S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

[31]J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

[32]C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

[33]W. England, M. S. Gordon, and K. Ruedenberg, Theoret. Chim. Acta (Berlin) **37**, 177 (1975).

[34]W. N. Lipscomb, Acc. Chem. Res. **6**, 257 (1973).

[35]D. A. Kleier, T. A. Halgren, J. H. Hall, Jr., and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).

[36]Localized-orbital procedures have been related via various approximations to one- and two-center energies; see for example W. England and M. S. Gordon, J. Am. Chem. Soc. **93**, 4649 (1971), and W. von Niessen, Theoret. Chim. Acta (Berlin) **38**, 9 (1975).

TABLE I. Detailed Multicenter Resolution of the SCF Energy[a] of H_2O

One-Center Terms

Center	Kinetic	E/N Attraction	E/E Repulsion	Total
H_1	0.37299	−0.55737	0.03821	−0.14617
H_2	0.37299	−0.55737	0.03821	−0.14617
O	74.23529	−176.46529	28.05106	−74.17894
Total	74.98126	−177.58003	28.12749	−74.47128

Two-Center Terms

Centers	Kinetic	E/N Attraction	E/E Repulsion	N/N Repulsion	Total
H_1 O	0.51636	−10.55520	4.77268	4.41902	−0.84714
H_2 O	0.51636	−10.55520	4.77268	4.41902	−0.84714
H_1 H_2	−0.00295	−0.23974	0.04992	0.34930	0.15653
Total	1.02976	−21.35013	9.59528	9.18733	−1.53776

Three-Center Terms

Centers	E/N Attraction	E/E Repulsion	Total
H_1 O H_2	−0.22049	0.18133	−0.03916
Total	−0.22049	0.18133	−0.03916

[a]In hartrees, computed at the experimental equilibrium structure, in the 6-311G(2d,p) basis. Total SCF Energy = −76.04820.

Table II. Basis Set Effect on Multicenter Resolution of the SCF Energy[a] in Homonuclear Diatomics

molecule	basis	$E^{(1)}$	$E^{(2)}$	E^{Total}
H_2	3-21G	−0.21918	−0.68457	−1.12293
	6-31G	−0.21973	−0.68728	−1.12674
	6-31G(d)	−0.21973	−0.68728	−1.12674
	6-311G(2d)	−0.21760	−0.69727	−1.13247
Li_2	3-21G	−7.21529	−0.33805	−14.76864
	6-31G	−7.33722	−0.19128	−14.86571
	6-31G(d)	−7.34151	−0.18338	−14.86639
	6-311G(2d)	−7.32979	−0.21057	−14.87014
N_2	3-21G	−53.29338	−1.71347	−108.30024
	6-31G	−53.83422	−1.19932	−108.86775
	6-31G(d)	−53.60852	−1.72559	−108.94262
	6-311G(2d)	−53.33633	−2.30342	−108.97607
F_2	3-21G	−98.62643	−0.39112	−197.64399
	6-31G	−99.22552	−0.19503	−198.64607
	6-31G(d)	−99.14409	−0.38496	−198.67314
	6-311G(2d)	−99.20595	−0.32654	−198.73844

[a] In hartrees, at the experimental equilibrium internuclear distances.

1 6

TABLE III. Convergence of the Multicenter Expansion of the SCF Energy for Polyatomic Molecules.

	$E^{(1)}$	$E^{(2)}$	$E^{(3)}$	$E^{(4)}$	E^{Total}
H_2O	−74.47128	−1.53776	−0.03916		−76.04820
N_2O	−180.06716	−4.06031	0.41399		−183.71348
CO_2	−183.56313	−4.58610	0.45692		−187.69231
NH_3	−53.78878	−2.43956	0.06536	−0.04855	−56.21153
NF_3	−350.98891	−1.95222	0.31382	−0.02189	−352.64920

[a]In hartrees, computed in the 6-311G(2d,p) basis at the experimental equilibrium structures.

TABLE IV. Energy Resolution[a] of the Reaction $2H + O \rightarrow H_2O$.

$2\Delta E^{(1)}_H$	443.8
$\Delta E^{(1)}_O$	393.1
$2\Delta E^{(2)}_{OH}$	−1063.2
$\Delta E^{(2)}_{HH}$	98.2
$\Delta E^{(3)}$	−24.6
ΔE^{Total}	−152.7
ΔE^{o}_0 (Exptl)[b]	−206.5

[a]In kcal/mole product, computed employing the experimental equilibrium structure of H_2O, in the 6-311G(2d,p) basis. The SCF energy of H in this basis is −0.49981 hartree and the energy of O is −74.80526 hartrees. [b] From the experimental atomization at 0 K, −219.532 ± 0.026 kcal/mol, plus the zero-point energy of H_2O, 12.869 kcal/mol. All experimental data from Ref. 19.

Table V. Multicenter Resolution of SCF Diatomic Reaction Energies[a]

$$H_2 + F_2 \rightarrow 2HF$$

$\Delta E^{(1)}_H$	130.76
$\Delta E^{(1)}_F$	70.41
$\Delta E^{(2)}$	-343.08
ΔE^{Total}	-141.92
ΔE^{o}_0 (Exptl)[b]	-134.5 ± 0.4

$$Li_2 + F_2 \rightarrow 2LiF$$

$\Delta E^{(1)}_{Li}$	90.25
$\Delta E^{(1)}_F$	98.97
$\Delta E^{(2)}$	-404.63
ΔE^{Total}	-215.41
ΔE^{o}_0 (Exptl)[b]	-215.1 ± 4.8

$$H_2 + Li_2 \rightarrow 2LiH$$

$\Delta E^{(1)}_H$	-151.46
$\Delta E^{(1)}_{Li}$	36.04
$\Delta E^{(2)}$	134.54
ΔE^{Total}	19.13
ΔE^{o}_0 (Exptl)[b]	18.6 ± 0.8

[a]In kcal/mole reactants, computed at the experimental equilibrium structures in the 6-311G(2d,p) basis. [b]Zero-point and thermal terms exploying experimental vibrational frequencies have been subtracted from the experimental energies. All experimental data from Ref. 19.

TABLE VI. Derivatives of Energy Components[a]

	$\partial E^{(1)}/\partial R$	$\partial E^{(2)}/\partial R$	$\partial E^{Total}/\partial R$
H_2	−0.219	0.227	0.008
Li_2	−0.056	0.049	−0.007
N_2	−1.740	1.939	0.198
F_2	0.745	−0.607	0.138
LiH	0.025	−0.028	−0.003
BH	−0.082	0.092	0.010
HF	−0.431	0.476	0.045
LiF	−0.445	0.453	0.008
BF	−0.850	0.883	0.033
CO	−0.314	0.438	0.124
	$\partial^2 E^{(1)}/\partial R^2$	$\partial^2 E^{(2)}/\partial R^2$	$\partial^2 E^{Total}/\partial R^2$
H_2	0.283	1.104	1.388
Li_2	−0.010	0.078	0.068
N_2	1.115	4.561	5.676
F_2	−2.289	3.440	1.151
LiH	−0.124	0.378	0.254
BH	−0.112	0.832	0.720
HF	0.576	1.709	2.284
LiF	1.939	−1.307	0.632
BF	1.187	0.698	1.S86
CO	−4.117	8.756	4.638

[a]Computed in the 6-311G(2d,p) basis at the experimental equilibrium internuclear distances. $\partial E/\partial R$ values are in hartrees/Å, $\partial^2 E/\partial R^2$ in hartrees/Å^2. The wavefunctions are allowed here to vary. If the orbital expansion coefficients are held constant, $\partial E^{(2)}/\partial R = \partial E^{Total}/\partial R$ and $\partial^2 E^{(2)}/\partial R^2 = \partial E^{Total}/\partial R^2$.

20

TABLE VII. First Derivatives of Energy Components[a] of H_2O

Derivatives With Respect to HOH Angle α	
$\partial E^{(1)}_H/\partial\alpha$	90.35
$\partial E^{(1)}_O/\partial\alpha$	480.71
$\partial E^{(2)}_{OH}/\partial\alpha$	−347.56
$\partial E^{(2)}_{HH}/\partial\alpha$	−3.88
$\partial E^{(3)}/\partial\alpha$	34.96
$\partial E^{Total}/\partial\alpha$	−2.62
Derivatives With Respect to Symmetric O–H Displacement R[b]	
$\partial E^{(1)}_H/\partial R$	−115.80
$\partial E^{(1)}_O/\partial R$	−598.83
$\partial E^{(2)}_{OH}/\partial R$	556.12
$\partial E^{(2)}_{HH}/\partial R$	−181.90
$\partial E^{(3)}/\partial R$	−54.75
$\partial E^{Total}/\partial R$	45.16

[a] In 10^{-3} hartree/radian and 10^{-3} hartree/Å, computed at the experimental equilibrium structure in the 6-311G(2d,p) basis. [b]The stretching symmetry coordinate $2^{-1/2}(\delta R_{OH1} + \delta R_{OH2})$ where R_{OH1} and R_{OH2} are the O–H internuclear distances. Note that all derivatives with respect to the third symmetry coordinate, the asymmetric stretch, are zero by symmetry.

TABLE VIII. Resolution of Second-Order Perturbation (MP2) Approximation to the Correlation Energy[a] in H_2O.

$E^{(1)}_H$	−0.02435
$E^{(1)}_O$	−0.30087
$E^{(2)}_{OH}$	0.06524
$E^{(2)}_{HH}$	−0.00021
$E^{(3)}$	−0.01448
E^{Total}	−0.23378

[a]In hartrees, computed at the experimental equilibrium structure in the 6-311G(2d,p) basis. Excitations from the 1s orbital of oxygen not included.

Accuracy in Ab Initio Reaction-Energy Computations. I. Compounds of First-Row Elements.

John R. Van Wazer,* Vladimír Kellö,† B. Andes Hess, Jr., and Carl S. Ewig
Department of Chemistry, Vanderbilt University, Nashville, Tennessee, 37235.

---

† Permanent address is Department of Physical Chemistry, Comenius University, 842 15 Bratislava, Czechoslovakia.

Abstract

Ab initio enthalpy computations were carried out for over forty gas-phase diamagnetic molecules (including eighteen hydrocarbons). All employed optimized geometries, basis sets ranging from 4-31G to 6-311++G(2df,2dp), and a series of electron-correlation approximations (MP2, MP3, MP4SDQ, and MP4SDTQ, as well as CCD, CCSD, CCSD+T(CCSD), and several CCSDT versions). The energies of forming the various molecules from the nuclei and electrons at zero K with no nuclear motion was calculated from experimental data and compared with the various ab initio values. The percentage difference between these experimental and ab initio values without correlation was found to be characteristic of each molecule regardless of the size of the basis set. For the hydrocarbons (and hydrides of other first-row atoms) these differences could be quantitatively related to the ratio of the number of hydrogens to the other atoms in each molecule. Electron correlation reduced this percentage difference by roughly a factor of two.

The enthalpies at 298 K of chemical reactions between the molecules were considered in terms of the disagreement between the experimental and theoretical enthalpies, with emphasis on generic classes of reactions, e.g. formation reactions involving (1) dehydrogenation of the common hydrides, or (2) combination of the homonuclear diatomics. Generic reactions showed up regularities in disagreements between experiment and theory. Reasons for large disagreements were probed.

A. Introduction

After years of slow progress, the computation of reaction enthalpies and other thermodynamic quantities by purely theoretical, nonempirical methods may soon emerge as a serious competitor not only to empirical and semiempirical approximations but also in many cases to direct experimental measurement. In the study reported herein, we investigate whether the common procedure of obtaining reaction enthalpies from differences in total molecular energies based on single-reference wavefunctions, with and without electron correlation, is a useful step in this direction. We also look into the assumption that a plethora of polarization functions and increasingly complicated and time-comsuming electron-correlation approximations should assure a small error in the theoretically computed enthalpy for a chemical reaction.

For SCF computations and alterations to them that incorporate electron correlation, $e.g.$ multiconfiguration SCF (MCSCF) and configuration interaction (CI), the variation theorem demands that a lower energy corresponds to greater accuracy. However, this does not apply to many correlation techniques, including the usual many-body perturbation procedure (Møller Plesset or MP) and the coupled-cluster method (CC), which can over- as well as under-estimate the correct total energy. Of course there is no a priori limit on the size or sign of energy *differences* among molecules. Although massive computations are thought to be required to achieve a close approach to the exact total energy of a molecule, it is reasonable to expect that, due to cancellation of errors, a rather poor but more easily computed approximation might be sufficient for obtaining consistently acceptable values for the very much smaller energies of chemical reactions. Since even Hartree-Fock limiting energies have been computed only for the atoms and a few diatomics, all meaningful reported ab initio reaction energies involving many-atom molecules are based inferentially on error cancellation.

The trick is to find an incomplete mathematical description for the molecular systems of interest in which error cancellation is routinely sufficient to allow trustworthy extension to the thermodynamics of chemical reactions for which data have not been obtained experimentally. One useful technique in such investigations is to employ increasing levels of mathematical description in order to see if the values of ΔH°_{298} and/or ΔG°_{298} smoothly approach acceptable values. Although many quantum-chemical papers report reaction energies without any evidence that they are not wildly in error, some studies have been concerned with the accuracy of ab initio energy computations.[1-3] These studies show unexpectedly large (and usually unexplained) energy errors for occasional reactions in an apparently well behaved series of examples.

In this contribution we look at the disagreements between the accepted experimental and theoretically computed enthalpies for the reactions of a variety of compounds of first-row atoms. This is done by utilizing two types of reactions for the formation of molecules: (1) from their constituent nuclei and electrons at 0 K without zero-point energy, and (2) from several general chemical reactions at room temperature . In addition, the results from various coupled-cluster electron-correlation schemes are intercompared and contrasted to the perturbation approximations. The purpose of this work is to characterize the behavior of individual compounds within the framework of Pople-type basis functions with many-body perturbation and coupled-cluster correlation corrections in order to find regularities that might serve to establish sets of mathematical descriptions, in conjunction with the choice of reactions and/or compounds, as being reliable for reaction-enthalpy computations.

B. Method

The modus operandi for our study of the enthalpies, ΔH°_{298}, of chemical reactions consists of the following operations: First the structural geometry of each molecule in a reaction must be optimized, using a particular mathematical description. From past experience we have found that gas-phase experimental geometries of the molecules involved in a reaction may not all be available or mutually consistent, so that, as a general rule, computed geometries ought to be employed. Note that geometries closer to good experimental values often result from optimizations involving the second-order many-body perturbation (MP2) correlation corrections[1,4] than from SCF optimizations alone. However for energy computations, it is more important that the geometries be consistent *with each other* rather than simply accurate to another decimal place. Next a computation[1,5] of its vibrational frequencies is needed for each molecule in order to account for nuclear motion, *i.e.* to convert from the computed theoretical energy at 0 K with no zero-point vibration to ΔH°_{298}. Fortunately neither this conversion factor nor the zero-point energy alone are very dependent on the choice of basis set used in the frequency computation (examples of which are given later).

All vibrations will be assumed here to be harmonic. Rotational energies are approximated by the high-temperature limit, *e.g.* 3RT/2 for polyatomics.

After the SCF-energy computations are completed for the chosen basis set(s) (which may be other[1] than that used in the geometrical optimization or the frequency computation) the various electron-correlation corrections are computed for each basis and geometry. Finally a simple program tabulates these data for each chosen chemical reaction, giving the experimental ΔH°_{298}, the ΔH°_{298} values corresponding to the various theoretical computations, and their disagreement, $\delta \Delta H^{\circ}_{298} = (\Delta H^{\circ}_{298})_{expd} - (\Delta H^{\circ}_{298})_{theor}$. Several such tabulations are shown later. Although $(\Delta H^{\circ}_{298})_{expd}$ values derived from

standard thermodynamic tables are usually accurate to no more than the first or second decimal place, they are occasionally presented herein to three decimal places to avoid rounding-off inconsistencies in Hess' law applications.

The main computer programs employed were CADPAC[5,6] and GAUSSIAN 86[7] for the SCF energies, geometries, frequencies and some of the many-body perturbation (MP) and the double-excitation coupled-cluster (CCD) correlation corrections, with COMENIUS[8,9] being used for all other electron correlations. The computations were run on our SCS 40/4 mini-supercomputer and the VAX 8800 cluster at the Vanderbilt Computer Center, with small jobs being relegated to the VAX 11/750 which serves as the entry port for the SCS. A few with large memory requirements were performed on the Cray X-MP/48 at the Pittsburgh Supercomputer Center. In all, over forty molecules were evaluated using as many as eleven different SCF computations (various basis sets and molecular geometries), each coupled with up to nine different correlation procedures per molecule. About a quarter of the molecules were subjected to nearly this full treatment and another quarter to less than four different SCF treatments, a few without any correlation.

Typical notation[10] for the mathematical model employed in a particular energy computation is exemplified by two examples: MP4SDQ/6-31G(d,p)//SCF/6-31G(d) and CCSD/6-31++G(2d,p)//MP2/6-31G(d,p). In the first example, the MP4SDQ refers to a fourth-order many-body perturbation correlation (MP4) including single (S), double (D), and quadruple (Q) excitations, with the triples (T) being omitted, as applied to an SCF computation in a 6-31G basis set. The d indicates that a sixfold d polarizing function was used for each first-row atom, and the p represents a threefold p polarizing function for each hydrogen. The information following the // describes how the molecular geometry was optimized (by SCF computations with a 6-31G(d) basis, *i.e.* no p's for the H's). The use of experimental geometry will hereinafter be denoted by //Exptl., while omission of the // and its following notation will indicate that the geometry was obtained by an SCF optimization in the same basis set as employed for the energy computation. In the second example, a coupled-cluster correlation covering single and double excitations was employed with a 6-31G basis enlarged by a diffuse sp set of functions on each first-row atom (the first +) and a similar s function added to each hydrogen (the second +). These diffuse functions were designed[11] to account for the expansion of an anionic molecule due to its negative charge. The types of polarization functions are explicitly shown, with (2d,p) standing for a pair of sixfold d's on each first-row atom and a threefold p on each hydrogen. [Note that the commonly used notations 6-31G* and 6-31G** are presented in this paper as 6-31G(d) and 6-31G(d,p).] The information following the // indicates that the geometrical optimization employed MP2 correlation with a smaller basis set, 6-31G(d,p).

Throughout this paper we shall use the term "mathematical description" as applied to a molecule or reaction to mean all the information conveyed by this notation plus, for values at temperatures above 0 K, the same kind of additional information for an accompanying frequency computation giving data which was not scaled or adjusted.

C. Reference Energies for Formation from Nuclei and Electrons

The overall reaction upon which any ab initio computed energy is implicitly based is, for a molecule $L_aM_bQ_c$,

$$aL^{+x} + bM^{+y} + cQ^{+z} + (ax+by+cz)e^- \rightarrow aL + bM + cQ \rightarrow L_aM_bQ_c. \qquad (1)$$
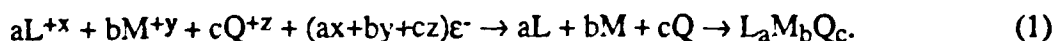
An experimental value for the energy of this reaction without nuclear motion (whereupon $\Delta H = \Delta E$) may be obtained from available data. Within experimental error, this quantity should represent the exact molecular energy, and hence it serves as a reference standard for any ab initio energy, E_{theor}. We shall symbolize this reference energy calculated for 0 K without zero-point energy as E_e, where the subscript e stands for "equilibrium," (the term employed in molecular-structure determinations to indicate internuclear angles and distances corresponding to the absence of nuclear motion). Note that E_e is uniquely suited as a reference for ab initio computations, since it permits examination of enthalpies of molecules individually, rather than the sum of all those appearing in a reaction.

The first step in calculating E_e by reaction (1) is to evaluate separately for each atom the sum of all of its experimental ionization potentials[12] to get the value (referenced to zero at infinite separation) for assembling that atom from its nuclei and electrons. Secondly the result is added to the tabulated atomic value[13] of ΔH°_0 for formation of the free gaseous atom to change the referencing to the element in its standard state. The third step consists of subtracting the sum of the resulting atomic enthalpy values times their stoichiometric proportions from the tabulated[13] heat of formation at 0 K of the molecule ($L_aM_bQ_c$) to obtain the value of ΔH°_0 for the formation of the molecule from its constituent nuclei and electrons at 0 K. To be directly comparable with the ab initio computed energies, the zero-point energy must be then be removed to reach the bottom of the potential well — the situation corresponding to no nuclear motion. The zero-point energy may be obtained from summing the experimental normal-mode vibrational frequencies of the molecule.[14] However, since such frequencies may be partially unavailable or suspect, the experimental zero-point energy has been approximated by multiplying our ab initio SCF/6–31G(d) zero-point energy by 0.900.

The resulting values of E_e (in hartrees) for twenty-five diverse gaseous molecules, listed in the second column of Table I, were checked for accuracy by calculating enthalpies

of various reactions among molecules at 298 K in the same way that these enthalpies were obtained from the ab initio data. (These accuracy test-values may be found in the footnotes to Tables VI through VIII in Section G and in the last column of Tables X and XI in Section H.) To summarize, the enthalpies based on E_e values consistently agreed with those obtained from heats of formation from the elements in their standard states, with the average of the absolute differences being less than 1 kcal/mol. For only 5 of 70 chemical equations was the absolute difference greater than 2 kcal per mole of the subject molecule, with the largest one being 2.7 kcal/mol. Thus, with no error being assigned to the conversion from 0 to 298 K, the values of E_e seem to be reliable to better than *ca.* 4×10^{-3} hartree for the first-row compounds of Table I.

The sum of the experimental ionization potentials of the constituent atoms represents 85% of the reference energy for the H_2 molecule and 99.7% to 99.99% for those molecules containing only first-row atoms with no hydrogen. Hydrogen-containing molecules based on one or more of the atoms C, N, O, or F exhibit summed ionization potentials ranging from 98.4% (for CH_4) through 99.5% (HCN) to 99.99% (HNO_3) of the respective molecular reference energy, E_e. Since such a large part of molecular energies are attributable to their atoms even when separated, knowledge concerning the first-row atoms may be extrapolated to their molecules. For these atoms the Hartree-Fock limit[15] is obtained quite accurately from numerical computations while its mass correction as well as the relativistic and Lamb corrections are calculable, so that an approximation to the correct value for the electron-correlation energy may be obtainable by difference.[15] These atomic computations indicate that SCF computations on the molecules of first-row atoms can at best represent no more than about 99.60% of the value of E_e and with perfect correlation no more than about about 99.95%

D. Analysis of Enthalpies in Terms of Relative Magnitudes

In comparing the various theoretical values computed for each molecule with its reference energy; we have found consistently that (excluding the H_2 molecule) for a given mathematical description *the difference*, $\Delta = E_e - E_{theor}$, *varies linearly with the reference energy*, E_e, while showing considerable scatter from molecule to molecule. This behavior is exemplified by Figure 1 in which the straight lines were obtained by a least-squares fit to the energies of the molecules listed in Table I, except for H_2. These lines correspond to the equation $\Delta = \alpha + \beta E_e$, for which α (a negative term) and β represent the intercept and slope respectively. Both of these terms in a give mathematical description have a consistently larger absolute value for SCF computations than for those with the various many-body-perturbation or coupled-cluster approximations. When the Δ-vs-E_e line for a set of SCF computations in a particular basis is contrasted with that from another SCF

computation in a larger basis, the latter is seen to exhibit a somewhat smaller slope, with the intercept term being essentially unchanged. This effect is demonstrated by Figure 2 in which the lines are fitted to the SCF data from our largest mathematical description, 6-311++G(2df,2pd)//MP2/6-31G(d,p), corresponding to 34 contracted functions for each first-row atom and 15 for each hydrogen, and to that from a much simpler description 6-31G//SCF/6-31G, with 9 and 2 contracted functions respectively. The pattern of scatter is about the same for the two bases. Indeed for SCF in all five of the basis sets employed, the deviation for a given molecule exhibits the same sign and about the same relative size. Graphs such as those of Figures 1 and 2 show that differences between the reference and theoretically computed energies for the formation from nuclei and electrons of individual gaseous molecules (without nuclear motion) may be logically intercompared in terms of a percentage disagreement, which includes a term for significantly large intercepts (α's) in the graphs in order that the less energetic molecules are not assigned inordinately high percentage deviations. Such a *percentage* disagreement, $100(\Delta-\alpha)/E_e$, equals $100 \cdot \beta$ where β is the slope of the graphed line.

Ab initio energies, E_{theor}, for twenty-five molecules in four basis sets are presented in Table I, which is ordered with respect to increasing E_e. Each energy was based on an MP2/6-31G(d,p) optimized geometry. In this table the theoretical data are reported in terms of an intercept-corrected percentage disagreement between the reference and theoretical values, $\%(\Delta-\alpha) = 100(E_e - E_{theor} - \alpha)/E_e$ with α being set to -0.1380 hartree for all SCF computations and to zero for all with correlation. An extra SCF data column for zero α is also given for the 6-31G(d,p) basis. The values of the energies as originally computed may be readily obtained to four decimal places from the $\%(\Delta-\alpha)$ values presented in Table I and can then be reinterpreted if desired.

For the twenty-four compounds of the first-row elements in the four basis sets shown in Table I, the values of $\%(\Delta-\alpha)$ lie in the range of 0.19 to 0.34% for the correlated energies ($\alpha = 0$) but are considerably larger, 0.33% to either 0.58% ($\alpha = -0.138$) or 0.79% ($\alpha = 0$), for the SCF. For H_2, these values are 0.15 to 0.76% with correlation ($\alpha = 0$) and -8.1 to -8.3% without ($\alpha = -0.138$). This latter range is large and negative due to $\alpha = -0.138$ being inappropriate since it was based on compounds of first-row atoms, not pure hydrogen. With $\alpha = 0$, $\%(\Delta-\alpha)$ for SCF for H_2 ranges from 3.4 to 4.0%, while it drops to the range of 0.04 to 0.19% with $\alpha = 0.040$. The average values given at the bottom of Table I represent the slopes ($100 \cdot \beta$) of the respective Δ-vs-E_e lines. Note that, if the the average were zero, the Δ-vs-E_e line would be horizontal indicating a fixed difference between E_e and E_{theor} equal to α; and, if its standard deviation were zero, there would be

no scatter in the points. With both zero, the ab initio values would precisely equal the reference ones.

The data of Table I show that these compounds form a set to which the H_2 molecule does not belong. Since the four hydrogens of methane, the most hydrogen-dominated molecule in this group of first-row compounds, contribute only about 5% of the reference energy of methane, the energetics of even this molecule is dominated by a first-row atom. Thus the H_2 molecule will be excluded from further discussion of Table I. Inspection of this table shows that the pattern of deviations from the average value for $\%(\Delta-\alpha)$ is about the same for the SCF in the various basis sets, with the C_2H_6, $(CH_3)_2O$, C_2, and C_2H_4 molecules exhibiting the larger positive deviations (0.11 to 0.06%); CH_2O, HCN, H_2O_2, CO_2, and N_2, the smaller ones (ca. 0.01 to –0.01%); and HF, H_2O, and F_2 having the larger negative ones (–0.11 to –0.06%).

Figure 3 depicts in two mathematical descriptions variations in the $\%(\Delta-\alpha)$ term as a function of E_e for the compounds (excluding H_2) listed in Table I. The points plotted from right to left in each figure correspond to reading down in the table. The horizontal lines at the left side of the figure correspond to the pertinent average values taken from the bottom of the table. The two sets of data at the top of each plot making up the figure demonstrate the effect of changing from $\alpha = 0$ to $\alpha = -0.138$ for the SCF data. This change lowers the points greatly on the right-hand side of the graph and only a small amount on the left, thereby causing the plot to lie more horizontally. The two data sets at the bottom of each plot show the effects of electron correlation. As previously noted, the pattern of the scatter of points for individual molecules does not change with basis set for the SCF case. But as exemplified by the CCSD and MP4SDTQ data of Figure 3, changes in the basis set do affect the pattern with correlation.

For many of the hydrocarbons, discussed in Section I below, only thermodynamic data for 298 K is available. Therefore a correction is needed to account for molecular rotation, translation, the contributions from excited vibrational states, and the conversion from enthalpy to energy. Ignoring the excited vibrational states, this correction equals $(^5/_2n-4)RT$, where n is the number of nuclei involved in equation (1) above. It is added to the energy summations of equation (1) while employing 298 K values of the heats of formation for the neutral atoms and the product molecule. A comparison of values of E_e based on reported $\Delta H^{\bullet}_0$ and $\Delta H^{\bullet}_{298}$ data for some representative compounds is given in Table II. It can be seen that the agreement between the two calculations is better than 1 kcal/mol of product.

The values of E_e and deviations from this value in two basis sets are shown for eighteen representative hydrocarbons in Table III. Because the tabulated heats of

formation[16] used for these molecules are not as well established as for the molecules shown in Table I, only four decimal places are reported for the values of E_e in Table III and α in $\%(\Delta-\alpha)$ is set to zero even for the SCF data. Hence the resulting values of $\%\Delta = 100(E_e - E_{theor})/E_e$ are given to only three significant figures. In this table the $\%\Delta$ values are seen to exhibit the same behavior noted previously for Table I and Figures 1–3 in that inclusion of correlation reduces the scatter of the $\%\Delta$ values from one molecule to another and lowers their average. However the values of E_e now appear in clusters which correspond to the number of carbon atoms per molecule, so that the kind of data treatment employed for the diverse group of molecules of Table I is no longer fruitful. Indeed a different approach is indicated since a homogeneous series of molecules such as the hydrocarbons offers a special opportunity for illuminating the role of molecular composition and structure (*vide infra*).

E. Analysis of Enthalpies in Terms of Molecular Compositions

After a lengthy and detailed examination of the hydrocarbon data, we concluded that the main unifying parameter is the ratio of the number of hydrogen to carbon atoms for each molecule, and that the data would be most readily interpretable using $\%\Delta$ directly (*i.e.* with zero α in all cases). Therefore Table III was organized with $\%\Delta = 100(E_e-E_{theor})/E_e$ listed in the order of increasing H/C ratio. In any of the seven data columns of the table, and within a group of molecules exhibiting a given H/C ratio, the values of $\%\Delta$ for either basis set are generally the same (within the reported errors of the thermodynamic data upon which the reference-energy values are based.[16]) Furthermore, with a few minor discrepancies, $\%\Delta$ is seen to increase proportionately to the H/C ratio for all of these 6–31G(d) and the 6–31G(d,p)//SCF/6–31G(d) data. This is shown graphically in Figure 4. By combining the equations for the straight lines shown in Figure 4 with the definition of $\%\Delta$, the following equations were obtained for the ratio of the reference to the theoretical energies:

$$E_e/E_{theor} = 1/\{\ 0.99335 + 0.00038(H/C)\} \qquad \text{for SCF/6–31G(d)} \qquad (2)$$
$$= 1/\{0.99662 + 0.00032(H/C)\} \qquad \text{for MP2/6–31G(d)} \qquad (3)$$
$$= 1/\{0.99689 + 0.00028(H/C)\} \qquad \text{for MP4/6–31G(d)} \qquad (4)$$
$$= 1/\{1.00000 + 0.00000(H/C)\} = 1 \qquad \text{for } E_{theor} = E_e$$

The fact that the same value of $\%\Delta = 0.739\%$ (with a standard deviation of only 0.0008) is found with SCF/6–31G(d) for ethylene, propene, either isomer of 2-butene, cyclopropane, and cyclobutane (for all of which H/C = 2) indicates that chemical structures in terms of bond orders and distances are not pertinent. Instead, the ratio of hydrogen to carbon atoms determines the effect. (This might also be inferred from the incongruously

large values of $\%\Delta$ found in every mathematical description for the H_2 molecule when compared to all compounds of the first-row elements, including the hydrocarbons.)

The kind of data interpretation employed above for the hydrocarbons may be extended to the general set of data reported in Table I. In doing this, we shall focus on the values given in Table I for SCF/6–31G(d,p) with $\alpha = 0.000$. Again for the hydrides of C, N, O, and F, $\%\Delta$ is found to vary quite linearly with the atomic ratio, H/X, of the hydrogen to the first-row atom, X, per molecule. These linear plots are shown in Figure 5. (Since the X=O plot of this figure was based on only two molecules (H_2O_2 and H_2O) and the same was true for the X=F plot (F_2 and HF), their linearity is not meaningful.)

The equations giving the linear dependence of $\%\Delta$ and of the equivalent E_e/E_{theor} ratio on the number of hydrogens per first-row atom for the individual atoms are presented below for SCF/6–31G(d,p).

For hydrides of carbon: $\%\Delta = 0.659 + 0.0326(H/C)$

$\qquad E_e/E_{theor} = 1/\{ \ 0.99341 + 0.000326(H/C)\}.$ (5)

For hydrides of nitrogen: $\%\Delta = 0.596 + 0.0311(H/N)$

$\qquad E_e/E_{theor} = 1/\{ \ 0.99404 + 0.000311(H/C)\}.$ (6)

For hydrides of oxygen: $\%\Delta = 0.553 + 0.0207(H/O)$

$\qquad E_e/E_{theor} = 1/\{ \ 0.99447 + 0.000207(H/O)\}.$ (7)

For hydrides of fluorine: $\%\Delta = 0.4985 + 0.0142(H/F)$

$\qquad E_e/E_{theor} = 1/\{ \ 0.99502 + 0.000142(H/F)\}.$ (8)

On the average, for any of the above:

$\%\Delta = [0.710 + 0.0532(5-N)] + [0.0386 + 0.0061(5-N)]](H/X),$ and

$E_e/E_{theor} = 1/\{[\ 0.99290 + 0.000532(5-N)] + [0.000386 + 0.000061(5-N)](H/X)\}$ (9)

where N is the atomic number of the first-row atom, B (not tested), C, N, O, or F.

It was found that this general linear equation, (9) above, encompassing the first-row atoms from atomic number 5 through 9 reproduces the data about as well as do the linear equations for the atoms treated individually.

Of the molecules listed in Table I that lie within the range of these equations, two do not fit well at all. They are the C_2 molecule for which $\%\Delta$ is 0.738 as compared to 0.659 from equation (5) and 0.657 from (9). The C_2 molecule is known[17] as being calculable to

an acceptable approximation only in a multiconfigurational description. The other problem ($\%\Delta = 0.585$, obs.; 0.553, eq. 5; and 0.550, eq. 9) occurs with ozone, O_3, for which a single-determinantal description has also been reported as inadequate.[18]

The H/X interpretive format developed for hydrocarbons has been extended not only to the hydrides of single kinds of atoms other than carbon but also to molecules based on several differing nonhydrogen atoms in the same molecule. The notation used in this approach is illustrated for three kinds of first-row atoms, $X_1, X_2, \& X_3$, for which an H:X_1:X_2:X_3 atomic ratio of 3:2:2:1 (as in fluoromethyl formate, FCH_2OOCH) would be reported as H/($2X_1,2X_2,X_3$) = 3/(2+2+1) = 0.60. Some data are depicted in Figure 5 by the lines for the H/(C,O) and H/(C,N) systems, as well as by the single points labeled as X= (2C,O), (C,F), and (C,2F) for various H/X atomic ratios. The one line based on three points for molecules containing H, X_1, and X_2 with a fixed proportion of X_1 to X_2 is shown in the figure labeled as X=(C,O). These three points (x) at H/X atomic ratio of 0, 1, and 2, which lie fairly well on a straight line, correspond to carbon monoxide, formaldehyde, and methanol respectively. The equations for $\%\Delta$ and E_e/E_{theor} for the H/(C,O) system in a SCF/6–31G(d,p) description are given below:

$$\%\Delta = 0.568 + 0.0331[H/(C,O)]$$
$$E_e/E_{theor} = 1/\{0.99432 + 0.000331[H/(C,O)]\} \tag{10}$$

For example, the rationale of equation (10) predicts that $\%\Delta$ for acetic acid and methyl formate will be essentially the same as that of monomeric formaldehyde [0.6055 (obs.) vs. 0.601 (eq.10)] and also the polymeric formadehydes.

When ratios (R = E_e/E_{theor}) rather than differences ($\Delta = E_e-E_{theor}$) are employed, E_e may be readily estimated from a theoretically computed energy obtained with a relatively small basis set (either with or without correlation). The E_e/E_{theor} relationships of equations 5 through 10 may be employed for this purpose, taking care to use the same mathematical description to compute both E_e and E_{theor}.

F. Factors Affecting Computed Enthalpies of Chemical Reactions

Before considering the computation of room-temperature heats of reaction, it seems desirable to clarify several pertinent details. Data for a few examples indicate that electron-correlation energies computed in a small 6-31G basis without excitation of the core electrons (i.e. using a frozen rather than a full core) is reduced from the core-inclusion value by about 1.5 kcal for ten-electron molecules to nearly 6 kcal for a forty-electron molecule (CF_4). For the 6-31++G(2d,p) basis, the reduction due to neglect of core-electron excitation is larger, 10.8 kcal for N_2 and 5.7 for NH_3. However, cancellations in

summing the contributions of the molecules involved in a chemical reaction usually lead to an overall effect on ΔH^o_{298} of the reaction of ca. 1 kcal/mol or less.

As exemplified in Table IV, the computed nuclear-motion correction to convert from an energy at 0 K with no nuclear motion to an enthalpy at 298 K changes by less than 2 (and often less than 1) kcal/mol between the various approximations in which the vibrational frequencies were computed. This is also true for the zero-point energy. This means that the choice of basis set to be employed for the vibrational-frequency computations becomes important only for accuracies better than about 1 kcal/mol. Although the zero-point energies obtained with a 6–31G basis set were scaled by multiplying by 0.9000 in the calculation of the essentially experimental number E_e, no scaling was employed in applying nuclear-motion corrections to obtain ΔH_{298} values from the ab initio data.

Considerable data were obtained for the effect on reaction-energy computations in the 6–31G(d,p) basis of using two different optimized geometries, based on SCF/6-31G(d) and MP2/6-31G(d,p) geometrical optimizations respectively. (Examples of such data may be seen in Tables VI–VIII in the next section.) For 163 chemical equations, the average absolute difference was 1.8 kcal/mol of product, with the absolute difference ranging from 0.01 to 3.4 kcal/mol. In most cases, the various correlation approximations exhibited differences that were about the same size for a particular equation, with the sign of the SCF differences usually being opposite to those with correlation.

The effect on the disagreement between the experimental and theoretical reaction enthalpies, $\delta\Delta H^o_{298} = (\Delta H^o_{298})_{explt} - (\Delta H^o_{298})_{theor}$, due to employing high-quality experimental geometries in place of computed ones was observed for a number of reactions, six of which are presented in Table V. Note that the effect is quite small, generally being less than 2 kcal/mol, so that the use of computed geometries are not the source of large disagreements.

G. Examples of Room-Temperature Intermolecular Reaction Enthalpies

Enthalpy tabulations for three chemical reactions have been selected from several hundred to illustrate our findings. In Table VI a number of theoretical approximations are presented for a reaction, $C_2H_6 + H_2 \rightarrow 2CH_4$, which exhibits consistently small values for the absolute value of the disagreement between experiment and theory, $|\delta\Delta H^o_{298}|$. This chemical equation is a particularly well-behaved example of the set of hydrocarbon reactions discussed below in Section J. For nineteen out of the thirty-nine entries in Table VI, the $|\delta\Delta H^o_{298}|$ values are smaller than 1.0 kcal and twenty-six are less than 2.0 kcal. Small but consistent correlation over-correction (negative values for $\delta\Delta H^o_{298}$ in this case) is seen for the basis sets without polarization functions on both the carbon and hydrogen atoms. Both positive and negative values of correlated $\delta\Delta H^o_{298}$ (all less than 1 kcal)

appear for the 6-31G(d,p) basis set with either optimized geometry, with nearly perfect agreement with experiment resulting from MP3, MP4SDQ, or CCSD correlation. Indeed the agreement is better than would be expected without elaborate attention to the smallest details. Note that adding diffuse s and p functions (indicated as ++) to the mathematical description reduces the efficacy of the correlation (other than MP2) in diminishing the absolute value of the disagreement between experiment and theory.

The results in Table VII, which is devoted to the $N_2 + 3H_2 \rightarrow 2NH_3$ reaction, are quite different. For the four smaller basis sets, in the upper half of the table, electron correlation causes an increase in $|\delta\Delta H^o_{298}|$ as compared to the SCF without correlation, so that most of the correlation "corrections" are in the wrong direction. However the addition of diffuse s and p functions (++) to the 6-31G(2d,p) basis set leads to three good agreements ($|\delta\Delta H^o_{298}| < 1$ kcal), with only one correlation going in the wrong direction [MP3/6-311++G(2d,p)//MP2/6-31G(d,p)]. It is interesting to note that for the largest basis set, 6-311G(2df,2pd) addition of diffuse s and p functions reduces the value of $|\delta\Delta H^o_{298}|$ for the SCF but does not cause it to decrease for the MP3 and MP4 correlation approximations.

For the reaction $CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$ shown in Table VIII we find that, for all basis sets having at least one polarization function on the first-row atoms, the SCF enthalpies are small and, with two exceptions, closer to the reported experimental value than with correlation. These exceptions are MP3/6-31G(d,p)//SCF/6-31G(d) and MP2/6-311G(2df,2pd)//MP2/6-31G(d,p). The former is only 1.8 kcal better than the SCF and the latter only 0.9 kcal. Not only is much of the correlation giving "corrections" in the wrong direction but the resulting errors are very large. Furthermore, inclusion of diffuse s and p functions into the 6-31G(d,p) basis set increases the values of $|\delta\Delta H^o_{298}|$ by about 25 kcal. Thus it is far from obvious what steps should be taken to obtain satisfactory theoretical results for this chemical reaction.

The data of Table VI are typical of a class of reactions which exhibit consistently modest disagreements between experiment and theory in small to large basis sets and for which the various correlation procedures consistently reduce the disagreements. The data of Table VII represent a different class, for which large disagreements and inept correlations in the smaller basis sets give way to improvements in both the SCF and the correlated values for the larger ones. (Note however that MP3 correlation overcorrects far too much in the large 6-311(2df,2pd) basis, either with or without diffuse (++) functions). Although demanding expensive computations for acceptable results, this behavior is clearly in accord with the principle of approaching limiting correct results as the size of the properly crafted basis set is increased. The data of Table VIII represents a class of

reactions in which correlation is quite uniformly deleterious, often ruinously so, and increasing the size of the basis set does not help much and often is harmful. The reaction of Table VII shows a pronounced drop in the absolute value of the disagreement upon adding diffuse functions, whereas the reaction of Table VIII responds oppositely. In running off tables such as these for hundreds of reactions resulting from combining various reactants chosen from the molecules of this study, we have observed multiple examples of each of these three kinds of behavior but were unable to codify these groupings usefully.

Tables VI–VIII also contain data for comparing the effect on $|\delta\Delta H^o_{298}|$ of molecular-geometry optimization in SCF/6–31G(d) as compared to MP2/6–31G(d,p). The differences between the two geometries with all else constant ranges from 0.2 to 5 kcal for the 6–31G(d,p) basis set in all three tables and from 1 to 3 kcal for the 6–31++G(2d,p) in Table VI. Also in all three tables, addition of triple excitations to either the MP4SDQ or CCSD correlations changes $\delta\Delta H^o_{298}$ by about the same amount for the perturbative as for the coupled-cluster method.

H. General Systematic Intermolecular Reaction Enthalpies

It would be desirable to set up a general "heat of formation" table that would involve reactions based only on diamagnetic gaseous reactants and products and could encompass all such molecules in their ground electronic and vibrational states. One possible group of reference molecules for the H/C/N/O/F compounds of this study consists of the common hydrides: H_2, CH_4, NH_3, H_2O, and HF. Any compound based on some combination of H, C, N, O, and/or F atoms can be formed from this group of hydrides, according to the following generic equation for "formation by dehydrogenation of the common hydrides," in which L and M stand for first-row elements.

$$a\,LH_x + b\,MH_z - \tfrac{1}{2}(xa+zb-c)\,H_2 \rightarrow L_aM_bH_c. \tag{11}$$
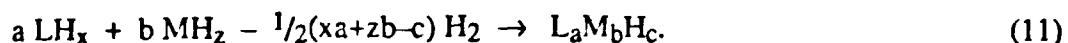
Since H_2, CH_4, NH_3, H_2O, and HF are the reference compounds for this scheme, their standard heats of formation must equal zero by definition. The experimental values of the heats of formation from them for 22 compounds are presented in kcal/mol in the second column, headed ΔH^o_{298}, of Table IX. The following paired columns of this table present the electron correlation method that gives the smallest absolute value for the disagreement between experiment and theory, and the resulting value of $|\delta\Delta H^o_{298}|_{min}$, for each mathematical description listed. Obviously the 6–31G basis is insufficient to give good ΔH_{298} values for these formation reactions, except for C_2H_6, CH_3NH_2, and CHF_3. The addition of polarization d functions is even less satisfactory for the disparate group of compounds studied, except for C_2, C_2H_2, C_2H_4, and CH_3NH_2. However, there is

considerable improvement when the hydrogen atoms are also given polarization p functions. Further improvement is achieved (for all but the reaction $2CH_4 - 4H_2 \rightarrow C_2$) by supplementing these polarization functions with diffuse functions on all of the atoms. Although there are only four examples for the 6-311G(2df,2pd)//MP2/6–31G(d,p) mathematical description, it does not seem to do as well as the preceding one. We thus infer that, at least for this collection of reactions in the basis sets employed, diffuse functions are more important for energy computations than are higher-order polarization functions. The last column of the table shows the fraction n/N, where n = the number of $|\delta\Delta H^o_{298}|$ values smaller than 1.1 kcal and N is the total number of these values considered for that particular reaction. When n is close to N, the enthalpy of the reaction presumably is easily computed theoretically.

Two other features of Table IX are worthy of discussion. First, an SCF entry in the table (showing that no correlation gives the best agreement with experiment) is obviously an indication that the basis set is inadequate for describing electron correlation. The 6–31++G(2d,p) basis appears to be appropriate and large enough to overcome this inadequacy for the subject reactions. Secondly, for none of the basis sets does a single correlation procedure give the minimum observed disagreement between experiment and theory for most of the reactions. Therefore in order to set up a usable theoretically derived thermodynamic table for a wide variety of molecules, we would have to compromise on a particular correlation to be used for all of the molecules — a correlation method that would be generally satisfactory. Table X explicitly shows the results for several possible compromises, particularly MP3/6-31G(d,p)//SCF/6-31G(d), MP4SDQ/6-31G(d,p)//SCF/-6-31G(d), (CCSD+T)/6-31++G(2d,p)//MP2/6-31G(d,p), and MP2/6-311G(2df,2pd)-//MP2/6–31G(d,p).

By applying Hess' law for energy additivity to any selected column in Table X, this reference scheme will give results analogous to that obtained from the usual compilations of heats of formation from the elements in their standard states. A convenient way of doing this is to subtract the values of the experimental ΔH^o_{298} (second column) for the reactants from those of the products and to do the same for the disagreement values, $\delta\Delta H^o_{298}$, from the chosen column. Subtracting the latter from the former difference gives the theoretically computed $\Delta H^{\bullet}_{298}$. Thus for the reaction $C_2 + H_2 \rightarrow C_2H_2$ (see Table XI), we find the experimental value of $\Delta H^{\bullet}_{298} = 89.95 - 235.96 - 0.00 = -146.01$ kcal and $\delta\Delta H^o_{298} = 0.04 - 0.79 - 0.00 = -0.75$ kcal for CCD/6-31G(d). $\delta\Delta H^o_{298} = -4.67 - (-5.25 + 0.00) = 0.58$ kcal for CCSD and similarly -16.81 for MP4SDTQ in the 6-31++G(2d,p)-//MP2/6–31G(d,p) approximation.

Another interesting set of reference standards consists of the stable, gaseous, diamagnetic, homonuclear-diatomic molecules, namely for this study: H_2, N_2, F_2 and perhaps C_2 even though it may be inadequately described by a single-determinant wavefunction. The stable form of O_2 is not diamagnetic and hence is omitted from this list. However, it can be substituted as the oxygen reference standard by any diamagnetic oxygen compound or even by the $^1\Sigma_g^+$ excited state of O_2, which is comparable to the ground states of the other diatomic elements used as standards. A tabulation similar to Table X but based on these diatomic molecules plus four different oxygen reference compounds – H_2O, O_3, CO, and CO_2 – is presented in Table XI. Before this latter table was started, a listing similar to Table IX was prepared of the smallest $\delta\Delta H^o_{298}$ values associated with formation from the diatomic elements, with water substituting for O_2.

Inspection of Table XI shows the 6–31G(d) basis set to be generally unsuitable (except for the four hydrocarbon molecules with CCD correlation), with the best results being obtained with the 6–31++G(2d,p) basis. Note that although the various oxygen-reference molecules may give quite different values for $\delta\Delta H_{298}$ for a particular molecule, all of the entries in a given column are tied together within the web of Hess' law. For example, for the redox equation $O_3 + CH_4 \rightarrow CO + 2H_2O$, the (CCSD+T)/6–31G(d,p)-//SCF/6–31G(d) description with oxygen-referencing to water gives $\delta\Delta H_{298} = (2.656 + 0.000) - (10.153 - 7.706) = 0.209$ kcal, as compared to $(0.000 - 5.311) - (2.188 - 7.706) = 0.207$ kcal with referencing to carbon monoxide. Using data from Table IX for this same reaction in the same mathematical description results in $\delta\Delta H_{298} = (10.362 + 0.000) - (10.153 - 0.000) = 0.209$ kcal. (Because of rounding errors, one more decimal place was employed above than appears in the tables, even though the experimental $\Delta_f H^o$ reported for ozone is significant to only the first decimal place.)

J. Hydrocarbon Reactions

From all of the above results it is clear that the hydrocarbons form a set of compounds that are in many ways distinct from the others. Therefore room-temperature enthalpies have also been investigated for reactions of the group of C_2 through C_4 hydrocarbons listed in Table III. Throughout this section we employ only molecular geometries computed in SCF/6-31G(d) and frequencies in SCF/6–31G. The values of $|\delta\Delta H^o_{298}|$ were found to be quite small for metathesis, isomerization, and hydrogenation reactions involving these hydrocarbons with each other and/or with hydrogen. Enthalpy data are presented in Table XII for seven of the nine possible independent chemical equations involving isomerizations between these compounds. The SCF data in this table exhibit fairly low values of $|\delta\Delta H^o_{298}|$. Addition of electron correlation gives appreciable improvement for only two reactions. The least energetic (regardless of sign) of the

reactions in this table is the cis/trans isomerization of 2-butene. In this case all of the $|\delta\Delta H^o{}_{298}|$ values with or without correlation are inappreciable, ranging between -0.52 and -0.63 kcal/mol. The two largest values of $|\Delta H^o{}_{298}|_{exptl}$ given in this table correspond to the butadiene-dicyclobutane and the butadiene-methylenecyclopropane pair. Not surprisingly these reactions also exhibit the larger $|\delta\Delta H^o{}_{298}|$ values. However, for the butadiene-dicyclobutane isomerization, the MP4SDTQ and CCD correlation corrections with the 6-31G(d) basis give $|\delta\Delta H^o{}_{298}|$ values of only 0.6 and 1.1 kcal.

In Table XIII the hydrocarbons are referenced to the common hydrides by means of the reaction $nCH_4 - (2n - {}^x/_2)H_2 \to C_nH_x$ (*i.e.* dehydrogenation of methane). This is analogous to the procedure used to construct Table X. Mathematical descriptions common to both tables are MP3 and CCD in the 6–31G(d) basis as well as MP3, MP4SDTQ, and CCSD+T(CCSD) at the 6–31G(d,p)//SCF/6–31G(d) level. (Note that the nuclear-motion corrections for the latter basis were derived from MP2/6-31G(d,p) frequency computations in Table X and from SCF/6-31G computations in Table XIII.) For each description, values of $|\delta\Delta H^o{}_{298}|$ for the hydrocarbons in both tables are generally smaller than for the other molecules of Table X. The $|\delta\Delta H^o{}_{298}|$ values in the SCF/6–31G(d) description for the C_2 and C_3 hydrocarbons lie in the range from -7 to -24 kcal/mol, so that the all-positive correlated $\delta\Delta H$ values for this basis correspond to modest overcorrections to the SCF.

Using the reaction ${}^n/_2C_2 + {}^x/_2H_2 \to C_nH_x$, the hydrocarbons are referenced in Table XIV to the homonuclear diatomics. (This is analogous to the analysis of the variety of compounds in Table XI.) Now by contrasting similar columns of Tables XI and XIV, we see the $|\delta\Delta H^o{}_{298}|$ values for the hydrocarbons to be quite similar to those of the other compounds, except for the CCD/6–31G(d) data which are consistently small for the hydrocarbons. These results, along with the $|\delta\Delta H^o{}_{298}|$ values of Table X, indicate that this mathematical description, which does not have excessive computational demands, is particularly well suited for the theoretical determination of enthalpies for reactions among hydrocarbons (including quite large ones).

The hydrocarbons were also referenced to the hydrogenation with dehydration of either carbon monoxide or carbon dioxide. The pertinent chemical equations were: $nCO + (n+{}^x/_2)H_2 - nH_2O \to C_nH_x$, or $nCO_2 + (2n+{}^x/_2))H_2 - 2nH_2O \to C_nH_x$. These two referencing schemes are compared graphically with those of Tables XIII and XIV in Figure 6. The use of CO as the carbon donor results in large values of $|\delta\Delta H^o{}_{298}|$, and CO_2 gives even larger ones (near 140 kcal/mol for the MP4SDTQ approximation). Of course all of these referencing schemes, whether they exhibit large or small values of $|\delta\Delta H^o{}_{298}|$, must give identical results (within rounding-off errors) when they are separately employed for computing $|\delta\Delta H^o{}_{298}|$ for any reaction in a particular mathematical description.

A glance at Figure 6 clearly demonstrates that the process of forming hydrocarbons by dehydrogenation of methane (the A curves) is consistently more accurate (having low values of $|\delta \Delta H^o_{298}|$) than the other formation processes (curves B–D). However for CCD correlation the process of formation from C_2 plus H_2 (curve B) also exhibits low-error behavior. In this figure, the stepwise shape of curves C and D, which consistently involve large values of $|\delta \Delta H^o_{298}|$, indicate that $|\delta \Delta H^o_{298}|$ is approximately proportional to the number of carbon atoms. For example, for curves C and D in the MP2 approximation, the fixed $\delta \Delta H^o_{298}$ contribution is respectively –17 and –30 kcal/mol/C atom. When the average disagreement between experiment and theory is small, as for the curves labeled A, this fixed contribution becomes insignificant.

K. Summary and Discussion

Data reduction — a time-consuming process which has played an important role in this work — has developed along two levels: The first is concerned with the fundamental thermodynamics of combining with electrons the point charges representing nuclei to form molecules, and the second with enthalpies of chemical reactions at room temperature. At both levels, the accuracies of the various quantum-chemical approximations were evaluated by comparison with experimentally based standards. An inherent difficulty lies in the fact that the experimental standards have their own errors which in some cases are as large as 5 or 10 kcal/mol. These errors result not only from problems in the experimental measurements but also from the processing and adjusting of the raw experimental data during the compilation of the reference thermodynamic tables. In addition they may be partly due to our manipulations in the case of the no-nuclear-motion reference standards, E_e's. At the end of Section C, the numbers (reported in the final columns of Tables X and XI) giving the results obtained from treating E_e values as if they were ab initio energies have been employed to place an upper limit on the errors of the E_e values. Of course these numbers also include the errors involved in converting energies involving no nuclear motion into ΔH^o_{298}, and the two sets of errors are not readily separable.

At the beginning of this work, we hoped to find some "magic" combination of basis set and correlation procedure that might lead to disagreements between theoretical and experimental room-temperature reaction enthalpies of no more than, say, 5 kcal/mol. This goal appears to have been reached in the limited case of hydrocarbons with the CCD/6-31G(d) mathematical description (see Tables VI, XII–XIV). However it does not seem very probable that a combination of any single-reference many-body-perturbation or coupled-cluster correlation procedure with even huge (by today's standards) basis sets will *consistently* give good accuracy for the ΔH_{298} of diverse reactions involving compounds of the first-row elements. Reactions between hydrocarbons, which may include the H_2

molecule, are quite well behaved in any mathematical description, as exemplified by the data of Table VI. The problem lies in reactions similar to those involving carbon dioxide and a hydrocarbon (see Table VIII and Figure 7) for which electron correlation often worsens rather than improves the theoretical reaction enthalpy, and also for which convergence with increasingly larger basis sets appears to be slow. This worsening is attributable to the subtraction of the sum of the correlation corrections for the reactants from the sum of those for the products — a process which can cause the resulting correlation to "correct" the wrong way.

This problem and others (including the fussy details treated in Section F) focused our attention on the theoretical energies obtained directly from ab initio computations and, for comparison purposes, on the estimation[19] from experimental data of the exact energy of formation of a molecule at zero K without zero-point energy from its nuclei and electrons. By casting the theoretical energy in the form of a percentage difference, $\%\Delta$, from the estimated exact energy for each molecule, it became immediately clear that the computed SCF energies as averaged over a group of molecules ranged from 99.2 to 99.6% of the exact energy depending on the basis set. Correlation caused an increase of 0.1–0.2 percentage units in these averages still leaving the MP4SDTQ/6–311++G(2df,2pd)// MP2/6-31G(d,p) description at only 99.8% of the ultimate energy limit.

In order to estimate what percentage of the exact energy corresponds to the Hartree-Fock limit for the molecules studied, we shall refer to the finding that the Hartree-Fock energies[15] for lone atoms represent the following percentages of the exact energies obtained from experimental ionization potentials[12]: 99.56% for C, 99.61 for N, 99.60 for O, and 99.61 for F. Thus, for these atoms the Hartree-Fock energy constitutes close to 99.60% of the exact energy (100.0% for H) and we think that this percentage ought to apply approximately to the molecules based on these atoms. The 6-311++G(2df,2pd)//MP2/6–31G(d,p) description gives SCF energies for the 25 molecules (excluding H_2) of Table I ranging from 99.24% of the exact value for CH_4, which is high in hydrogen, to 99.54% for F_2, with an average value of 99.41% for all 25 molecules. This average over the molecules is about 0.2 percentage point above the percentage corresponding to the Hartree-Fock limit obtained from averaging over the subject atoms, while the methane value is 0.3 point above the limit for C and the F_2 value is 0.1 point above the F limit. Not surprisingly it appears that the molecular composition (with emphasis on hydrogen) seems to affect the closeness to the Hartree-Fock limit of a molecular SCF energy obtained in a given basis set.

The sign and approximate size of the divergence from the average $\%\Delta$ for a particular molecule seems to be characterized by SCF computations on that molecule (regardless of basis set). An investigation of this behavior for the varied group of

hydrocarbons showed that the error of the SCF computations (as measured by the ratio of the exact to the SCF energy with no nuclear motion) was linearly related to its atomic ratio of hydrogen to carbon, H/C, and not to its bond structure. (The only all-hydrogen diamagnetic molecule, H_2, consistently exhibits a manyfold larger value of $\%\Delta$ than does any molecule based on first-row atoms.) Thus in effect the hydrogen nuclei may be considered as interlopers in a rather foreign electronic environment. The concept of interloper hydrogen has been applied to hydrides of the first-row atoms other than carbon and also has been roughly fitted to molecules containing several kinds of first-row atoms. Although the emphasis is on hydrogens, the intercept terms of the linear equations relating the SCF error to the hydrogen/other-atoms ratio deal with the errors of hydrogen-free molecules as well.

Since the many-body-perturbation or coupled-cluster correlation procedures act to minimize the total energy through excitations to virtual orbitals made available by the selected basis set, both the number and nature of excitations will differ from one basis set to another. Considerable weight may thus be given to excitations that would not contribute significantly to the chosen correlation at the Hartree-Fock limit. Because of this, correlation is likely to mask the pattern of errors in the SCF energies (as in Sect. D above).

For the lone first-row atoms the sum of the relativistic, the Lamb, and the mass corrections[15] represents 0.037 (C), 0.049 (N), 0.064 (O), and 0.080 (F) percent of their exact energies so that the Hartree-Fock limit with perfect electron correlation corresponds to what is left (averaging about 99.943% of E_e). Presumably this energy limit including correlation may also apply to the molecules. The MP4SDTQ/6–311++G(2df,2pd)-//MP2//6-31G(d,p) description corresponds to 99.805% of the exact energy for the NH_3 and N_2 molecules, in comparison to a value of 99.951% for the energy limit with correlation for the nitrogen atom. Thus the energies of these two molecules in this description seem to be about 0.14 percentage points above their limiting values. Unfortunately these differences are of course still quite large on the scale of typical chemical reactions.

The concepts of "isodesmic[20]" reactions as applied to errors in quantum-mechanically computed molecular energies were developed for much more primitive computrational procedures than the ones employed in this paper and are now probably obsolete. These concepts seem to have been applicable to chemical reactions between reasonably similar molecules; but they could have been simply a complicated way of saying that errors in small numbers are apt to be smaller than those in large numbers. In any case, our approach to finding simple mathematical relationships for the errors, an appreoach

which led to the "interloper-hydrogen" concept, indicates that a rationale based on bond structures is probably inappropriate.

The energy, E_e, as developed herein seems to us to be particularly well suited for the quantitative exploration of the detailed electronic structure of molecules and an updating of the presently generally accepted notions of the way atoms are incorporated into molecules. The absence of vibrational energies and the implicit full description of the subject molecule should assure that this limiting energy will properly reflect the electronic structure. By using the H/X ratio to estimate this energy from that of an SCF computation in a modest basis set, values may be found for molecules for which the proper experimental thermodynamic data are not available or which have too many atoms for computations in a large basis set with good correlation. Such an estimated energy may be combined with others and then converted to a desired but otherwise unavailable heat of reaction. This approach may be considered as a semi-empirical method in which the extrapolations based on experimental data take place following a lower-level ab initio compuatation rather than to establish shortcuts to be employed in the quantum computation itself. Additional work will be needed to test further and to develop suitable accuracy for this suggested procedure for estimating reaction enthalpies at specific temperatures.

Acknowledgment

Literature References

(1)   Hehre, W. J.; Radom, L.; Schleyer, P. v.R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.

(1)(b) For an unusually detailed analysis of various ab initio methods, see Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.*, **1984**, *80*, 3265

(2)   Dewar, M. J. S.; Storch, D. M. *J. Am. Chem. Soc.*, **1985**, *107*, 3898.

(3)   Kello, V. *Chem. Zvesti,,* **1982**, *36*, 289.

(4)   DeFrees, D. J.; Krishnan, R.; Schlegel, Pople, J. A. *J. Am. Chem. Soc.*, **1982**, *104*, 5576.

(5)   The CADPAC program gives both the energy gradients and the force-constant matrices analytically in the MP2 approximation, as described by Simandiras, E. D.; Handy, N. C.; Amos, R. D. *Chem. Phys. Letters* **1987**, *133*, 324.

(6)   Amos, R. D.; Rice, J. E. *CADPAC: The Cambridge Analytical Derivatives Package*, issue 4.0, Cambridge University, 1987.

(7)   Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Raghavachari, K.; Melius, C. F.; Martin, R. L.; Stewart, J. J. P.; Bobrowicz, F. W.; Rohlfing, C. M.; Kahn, L. R.; Defrees, D. J.; Seeger, R.; Whiteside, R. A.; Fox, D. J.; Fleuder, E. M.; Pople, J. A. *GAUSSIAN 86;* Carnegie-Mellon Quantum Chemistry Publishing Unit: Pittsburgh, PA, 1984.

(8)   Noga, J.; Kello, V.; Urban, M. *Program COMENIUS,* Version 1987, Initial comments, 1987.

(9)   Urban, M.; Cernusak, I.; Kello, V.; Noga, J. in *Computational Methods in Chemistry*, Vol. I. (editor: S. Wilson) Plenum, New York, 1987. Also see Carsky, P.; Urban, M. *Ab Initio Calculations. Methods and Applications in Chemistry. Lecture Notes in Chemistry.* Vol, 16, Springer, Berlin, 1980.

(10)   See ref. 1 for additional details. In this paper we use the terms 6–31G(d) and 6–31G(d,p) for the basis sets usually designated as 6–31G* and 6–31G**.

(11)   Clark. T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P.v. R. *J. Comp. Chem.* **1983**, *4*, 294.

(12)   Moore, C. E. *Ionization Potentials and Ionization Limits Derived from the Analysis of Optical Spectra, NSRDS-NBS-34,* Natl. Bur. Stands., Washington, D.C., 1970.

(13)   Wagman, D. D.; Evans, W. H.; Parker, V. B.; Halow, I.; Bailey, S. M.; Schumm, R. H. *Selected Values of Chemical Thermodynamic Properties, NBS Technical Note 270-3,* Natl. Bur. Stands., Washington D. C., 1968. Also see Stull, D. R.; Prophet, H. (eds.) *JANAF Thermochemical Tables, Natl. Bur. Stands.,* Washington. D.C., 1971; and

Thermodynamic Properties of Individual Substances (in Russian); Glyshko, W. P., Ed.; Science: Moscow, 1982.

(14) Shimanouchi, T. *J. Phys. Chem. Ref. Data* 1977, 6, 993.

(15) Viellard, A.; Clementi, E. *J. Chem. Phys.* 1968, 49, 2415, and references therein.

(16) Pedley, J. B.; Rylance, J., *CATCH Tables*, University of Sussex, 1977. Also see Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*, Academic Press, London, 1970, which contains much of the data later reported by Pedley and Rylance. Except for our use of the value of $\Delta_f H^\circ$ = +66.2 kcal reported by Cox and Pilcher for cyclopropene, we employ the *CATCH* data exclusively for the intercomparisons of hydrocarbons. The use of the *CATCH* value of 42.3±0.6 kcal for cyclopropane led to %Δ values for this molecule in Table III that were consistently much too high with respect to the other molecules and corresponded to an H/C atomic ratio for this molecule of *ca.* 2.5+ instead of the expected 1.33. Use of the $\Delta_f H^\circ$ of Cox and Pilcher brought the %Δ values for cyclopropene down to the expected range. Generally the two compilations agreed to better than 0.1–0.2 kcal..

(17) Rohlfing, C. M.; Martin, R. L. *J. Phys. Chem.* 1986, 90, 2043.

(18) An overview of previous calculations on ozone, especially those employing configuration mixing, is given by Wilson, C. W., Jr.; Hopper, D. G. *J. Chem. Phys.* 1981, 74, 595. See in particular Goddard, W. A., III; Dunning, T. H., Jr.; Hunt, W. J.; Hay, P. J. *Acc. Chem. Res.* 1973, 6, 368.

(19) Absar, I.; Schaad, L. J.; Van Wazer, J. R. *Theoret. Chim. Acta* 1973, 29, 173.

(20) Hehre, W. E. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* 1970, 92, 4796.

Table I

Intercept-Corrected Percentage Difference between Reference and Theory for Energies of Formation from Nuclei and Electrons of Various Gaseous Molecules at 0 K without Zero-Point Energy

Percentage Difference, $\%(\Delta-\alpha) = 100(E_e - E_{theor} - \alpha)/E_e$

Compd.	Reference Energy (hartrees)	α = 0.000	6-31G(d,p) SCF (-0.138)	6-31G(d,p) MP4SDTQ (0.000)	6-31G(d,p) CCSD (0.000)	6-31++G(2d,p) SCF (-0.138)	6-31++G(2d,p) MP4SDTQ (0.000)	6-31++G(2d,p) CCSD (0.000)	6-311G(2df,2pd) SCF (-0.138)	6-311G(2df,2pd) MP4SDTQ (0.000)	6-311++G(2df,2pd) SCF (-0.138)	6-311++G(2df,2pd) MP4SDTQ (0.000)
H2	-1.17353	3.5966	-8.1634	0.7635	0.7149	-8.1694	0.7533	0.7039	-8.3074	0.1559	-8.3083	0.1551
CH4	-40.52218	0.7909	0.4503	0.3296	0.3363	0.4447	0.2088	0.3075	0.4229	0.2111	0.4229	
NH3	-56.58497	0.6888	0.4449	0.3244	0.3295	0.4311	0.2795	0.2883	0.4024	0.1961	0.4034	0.1955
C2	-75.93921	0.7384	0.5566	0.2696	0.3205	0.5486	0.2452	0.3010	0.5293	0.1923	0.5303	
H2O	-76.47724	0.5941	0.4137	0.3215	0.3249	0.3974	0.2648	0.2725	0.3672	0.1880	0.3684	
C2H2	-77.35559	0.6928	0.5144	0.3186	0.3338	0.5036	0.2871	0.3053			0.4798	
C2H4	-78.60567	0.7216	0.5461	0.3205	0.3308	0.5370	0.2909	0.3039	0.5146		0.5153	
C2H6	-79.84156	0.7557	0.5828	0.3225	0.3310	0.5782	0.2956	0.3063			0.5577	
HCN	-93.46337	0.6326	0.4849	0.2943	0.3106	0.4779	0.2663	0.2857			0.4533	
HF	-100.52660	0.5127	0.3755	0.3235	0.3259	0.3595	0.2635	0.2691	0.3384		0.3313	
N2	-109.58813	0.5956	0.4697	0.2876	0.3039	0.4622	0.2593	0.2785	0.4255	0.1985	0.4368	0.1953
N2H4	-111.92054	0.6594	0.5361	0.3157	0.3225	0.5253	0.2770				0.5215	
CO	-113.37565	0.5655	0.4438	0.2931	0.3075	0.4342	0.2552	0.2724	0.4048	0.1965	0.4095	
CH2O	-114.56103	0.6055	0.4850	0.3074	0.3175	0.4754	0.2690	0.2824			0.4494	
CH3OH	-115.77728	0.6318	0.5126	0.3157	0.3218	0.5032	0.2740	0.2836			0.4762	
CH3F	-139.82863	0.5645	0.4658	0.3162	0.3214	0.4332	0.2372				0.4289	
H2O2	-151.64219	0.5734	0.4824	0.3090	0.3148	0.4735	0.2631	0.2732			0.4443	
(CH3)2O	-155.08559	0.6528	0.5638	0.3131	0.3196						0.5529	
N2O	-184.76111	0.5943	0.5196	0.2879	0.3122	0.4885	0.2175	0.2175			0.4836	
CO2	-188.68981	0.5625	0.4894	0.2960	0.3112	0.4571	0.2211		0.4483	0.1924	0.4527	
F2	-199.66812	0.4985	0.4294	0.3090	0.3137	0.4246	0.2645	0.2722			0.3948	
O3	-225.55647	0.5845	0.5233	0.2959	0.3167	0.4897	0.2131		0.5032	0.2544	0.4863	
CH2F2	-239.15450	0.5251	0.4674	0.3120	0.3174	0.4600	0.2794				0.4560	
OF2	-274.86288	0.5160	0.4658	0.3064	0.3130	0.4323	0.2165				0.4302	
HNO3	-281.05214	0.5739	0.5248	0.3000		0.5155					0.5110	
Average (excluding H2)		*0.618*	*0.490*	*0.308*	*0.320*	*0.472*	*0.261*	*0.282*	*0.436*	*0.204*	*0.458*	*0.195*
Std.Deviation (··)		*0.084*	*0.050*	*0.015*	*0.009*	*0.051*	*0.026*	*0.022*	*0.063*	*0.022*	*0.057*	*0.000*

Table II.
Reference Energies for Formation from Nuclei and
Electrons without Nuclear Motion
Calculated From Both ΔH°_0 and ΔH°_{298} Data

| Compd. | Values of E_e Calculated from | | |
	ΔH°_0 Data (hartrees)	Difference (kcal/mol)	ΔH°_{298} Data (hartrees)
H_2	−1.17353	0.35	−1.17408
CH_4	−40.52218	0.06	−40.52228
NH_3	−56.58497	0.01	−56.58495
C_2	−75.93921	0.0	−75.9392
CO	−113.37565	0.50	−113.37645
CO_2	−188.68981	0.46	−188.69055
H_2O	−76.47724	0.14	−76.47746
C_2H_2	−77.35559	0.13	−77.35580
C_2H_4	−78.60567	0.01	−78.60568
C_2H_6	−79.84156	−0.33	−79.84103

Table III.

Percentage Disagreement[a], %Δ, Between Reference and Theory for Energies of Formation from Nuclei and Electrons of Gaseous Hydrocarbons at 0 K without Zero-Point Energy

Compound	H/C Ratio	Reference Energy (hartrees)[b]	Percentage Disagreement,[a] $\%\Delta = 100(E_e - E_{theor})/E_e$						
			SCF	MP2	MP3	MP4SDQ	MP4SDTQ	CC[c]	CCSD+T[d]
6-31G(d) basis set									
Butadiyne, C_4H_2	1/2	−153.5239	0.668	0.350	0.344		0.321	0.342	
Acetylene, C_2H_2	1	−77.3553	0.695	0.376	0.362		0.342	0.359	
Allene, C_3H_4	1 1/3	−116.6828	0.704	0.385	0.364		0.346	0.361	
Propyne, C_3H_4	1 1/3	−116.6855	0.704	0.382	0.366		0.346	0.362	
Cyclopropene, C_3H_4	1 1/3	−116.6505	0.709	0.383	0.365		0.348	0.362	
Methylenecyclopropane	1 1/2	−155.9943	0.710	0.384	0.364		0.346	0.361	
Butadiene, C_4H_6	1 1/2	−156.0285	0.711	0.389	0.368		0.351	0.365	
Bicyclobutane, C_4H_6	1 1/2	−155.9888	0.716	0.385	0.367		0.350	0.365	
Cyclobutane, C_4H_8	2	−157.2445	0.729	0.399	0.375		0.359	0.371	
trans-2-Butene, C_4H_8	2	−157.2571	0.729	0.402	0.377		0.361	0.373	
cis-2-Butene, C_4H_8	2	−157.2558	0.730	0.403	0.378		0.361	0.374	
Cyclopropane, C_3H_6	2	−117.9210	0.731	0.401	0.377		0.362	0.374	
Propene, C_3H_6	2	−117.9319	0.730	0.405	0.379		0.363	0.375	
Ethylene, C_2H_4	2	−78.6057	0.730	0.409	0.382		0.365	0.377	
Butane, C_4H_{10}	2 1/2	−158.4842	0.748	0.416	0.387		0.372	0.383	
Propane, C_3H_8	2 2/3	−119.1625	0.754	0.422	0.392		0.377	0.387	
Ethane, C_2H_6	3	−79.8407	0.767	0.434	0.401		0.386	0.395	
Methane, CH_4	4	−40.5222	0.807	0.468	0.429		0.414	0.420	
Average for all molecules			*0.726*	*0.400*	*0.376*		*0.359*	*0.373*	
Standard Deviation			*0.030*	*0.026*	*0.018*		*0.020*	*0.017*	
6-31G(d,p)//SCF/6-31G(d) basis set									
Butadiyne, C_4H_2	1/2	−153.5239	0.666	0.342	0.335	0.330	0.312	0.329	0.313
Acetylene, C_2H_2	1	−77.3553	0.690	0.357	0.343	0.337	0.322	0.336	0.321
Allene, C_3H_4	1 1/3	−116.6828	0.698	0.358	0.336	0.331	0.317	0.330	0.315
Propyne, C_3H_4	1 1/3	−116.6855	0.698	0.356	0.338	0.333	0.319	0.332	0.318
Cyclopropene, C_3H_4	1 1/3	−116.6505	0.703	0.355	0.335	0.332	0.318	0.331	0.317
Bicyclobutane, C_4H_6	1 1/2	−155.9888	0.692	0.341	0.318	0.315	0.302	0.315	0.301
Cyclopropane, C_3H_6	2	−117.9210	0.723	0.360	0.333	0.330	0.318	0.329	0.317
Propene, C_3H_6	2	−117.9319	0.721	0.363	0.335	0.331	0.319	0.330	0.317
Ethylene, C_2H_4	2	−78.6057	0.721	0.368	0.338	0.333	0.321	0.332	0.319
Propane, C_3H_8	2 2/3	−119.1625	0.744	0.367	0.333	0.329	0.319	0.329	0.317
Ethane, C_2H_6	3	−79.8407	0.755	0.372	0.336	0.331	0.322	0.330	0.320
Methane, CH_4	4	−40.5222	0.791	0.389	0.344	0.337	0.330	0.336	0.327
Average for all molecules			*0.717*	*0.361*	*0.335*	*0.331*	*0.318*	*0.330*	*0.317*
Standard Deviation			*0.034*	*0.013*	*0.006*	*0.006*	*0.007*	*0.005*	*0.006*

[a] Columns 4 through 10 show values of $\%\Delta = 100\,(\Delta-\alpha)/E_c$, for α = zero.   [b] Calculated using $\Delta_f H°$ for 298.15 K (based on data from CATCH Tables[16]).   [c] CCD for 6-31G(d) and CCSD for 6-31G(d,p).   [d] CCSD+T signifies CCSD+T(CCSD)...

Table IV.
Computed Nuclear-Motion Energy Correction
and Zero-Point Energy (in italics) for Various Molecules
(kcal/mol)

Molecule	Calculation Employed			
	4–31G	6–31G	MP2/6–31G(d,p)	6–31++G(2d,p) //MP2/6–31G(d,p)
H_2	8.123	8.123	8.071	8.089
	6.641	*6.642*	*6.590*	*6.608*
C_2	4.154	4.130	4.598	4.193
	2.648	*2.648*	*3.117*	*2.711*
CH_4	31.860	31.867	31.045	31.305
	30.071	*30.078*	*29.251*	*29.513*
C_2H_4	36.463	36.439	34.797	35.861
	34.581	*34.581*	*32.894*	*33.989*
NH_3	24.652	24.717	24.077	24.794
	22.780	*22.835*	*22.283*	*23.001*
H_2O	15.804	15.890	15.521	16.293
	14.025	*14.111*	*13.743*	*14.514*
CO	4.779	4.749	4.519	4.984
	3.297	*3.268*	*3.038*	*3.467*
CO_2	8.965	8.934	8.917	9.621
	7.307	*7.284*	*7.256*	*8.033*

Table V.

Comparison of the Effect on $\delta\Delta H°_{298}$ of Employing Experimental Instead of Computed Geometries.

Compound C_nH_x	Correlation Method	$\delta\Delta H°_{298}$ (kcal/mol)	
		Exptl. Geom.	Calc. Geom.
Generic Reaction: $nCH_4 - (2n - x/2)H_2 \rightarrow C_nH_x$			
C_2H_2	None	−13.03	−12.75
	MP2	3.38	2.33
	MP3	-0.92	-1.56
C_2H_4	None	−7.50	−7.16
	MP2	1.10	0.28
	MP3	0.55	1.13
C_2H_6	None	−4.30	−3.03
	MP2	1.62	2.30
	MP3	1.19	1.79
Generic Reaction: $nCO + (n+x/2)H_2 - nH_2O \rightarrow C_nH_x$			
C_2H_2	None	−11.61	−11.77
	MP2	−18.58	−17.73
	MP3	−13.50	-13.13
C_2H_4	None	−36.25	−36.32
	MP2	−33.77	−33.12
	MP3	−27.93	−27.83
C_2H_6	None	−30.71	−30.70
	MP2	−36.06	−34.69
	MP3	−26.46	−25.98

Table VI.
Comparison of Various Levels of Approximation for ΔH°_{298} (in kcal) for the Gaseous Reaction $C_2H_6 + H_2 \rightarrow 2CH_4$.

Correlation[a]	ΔH°_{298}	$\delta\Delta H^{\circ}_{298}$[b]
4–31G basis, with vib. approx. in SCF/4–31G		
SCF	-19.81	4.29
MP2(full)	-14.79	-0.73
6–31G basis , with vib. approx. in SCF/6–31G		
SCF	-19.77	4.25
MP2(full)	-14.61	-0.91
MP2	-14.64	-0.88
MP3	-14.42	-1.10
6–31G(d) basis, with vib. approx. in SCF/6–31G		
SCF	-18.55	3.03
MP2	-13.22	-2.30
MP3	-13.73	-1.79
MP4SDTQ	-13.01	-2.51
CCD	-13.71	-1.81
6-31G(d)//Exptl., with vib. approx. in SCF/6–31G		
SCF	-19.82	4.30
MP2	-13.91	-1.62
MP3	-14.33	-1.19
6–31G(d,p)//SCF/6–31G(d), vibs. in MP2/6–31G(d,p)		
SCF	-18.17	2.65
MP2	-14.65	-0.89
MP3	-15.62	0.10
MP4SDQ	-15.64	0.19
MP4SDTQ	-14.92	-0.60
CCSD	-15.46	-0.06
CCSD+T(CCSD)	-14.85	-0.67
6–31G(d,p)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	-18.19	3.67
MP2	-14.62	-0.90
MP3	-15.60	0.08
MP4SDQ	-15.60	0.08
MP4SDTQ	-14.88	-0.64
CCSD	-15.42	-0.10
CCSD+T(CCSD)	-14.81	-0.71
CCSDT1	-14.77	-0.75
CCSDT2	-14.88	-0.64
CCSDT3	-14.87	-0.65
6–31++G(2d,p)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	-18.68	3.16
MP2	-16.33	0.82
MP3	-17.91	2.39
MP4SDQ	-17.82	2.30
MP4SDTQ	-17.01	1.49
CCSD	-17.55	2.04
CCSD+T(CCSD)	-16.46	1.39
6–311++G(2df,2pd)//MP2/6–31G(d,p), vibs. MP2/6–31G(d,p)		
SCF	-18.46	2.94

[a] Including only valence excitations (frozen core), except for those denoted as (full).
[b] $\delta\Delta H^{\circ}_{298} = (\Delta H^{\circ}_{298})_{expl} - (\Delta H^{\circ}_{298})_{calc}$, with $(\Delta H^{\circ}_{298})_{expl} = -15.52$ kcal and with ΔH°_{298} from E_e values at 0 K without zero-point energies = -15.30 kcal.

Table VII.
Comparison of Various Levels of Approximation for
ΔH°_{298} (in kcal) for the Gaseous Reaction
$N_2 + 3H_2 \rightarrow 2NH_3$

Correlation	ΔH°_{298}	$\delta \Delta H^{\circ}_{298}$[a]
6–31G basis, with vib. approx. in SCF/6–31G		
SCF	–33.21	11.17
MP2	3.62	–25.66
MP3	–10.03	–12.01
6–31G(d) basis, with vib. approx. in SCF/6–31G		
SCF	–9.19	–12.85
MP2	1.70	–23.74
MP3	–5.14	–16.90
MP4SDTQ	4.76	–26.80
CCD	–1.93	–20.11
6–31G(d,p)//SCF/6–31G(d) basis, vibs. in MP2/6–31G(d,p)		
SCF	–15.18	–6.86
MP2	–9.93	–12.11
MP3	–17.46	–4.58
MP4SDQ	–12.44	–9.60
MP4SDTQ	–8.17	–13.87
CCSD	–12.59	–9.45
CCSD+T(CCSD)	–9.41	–12.63
6–31G(d,p)//MP2/6–31G(d,p) basis, vibs. in MP2/6–31G(d,p)		
SCF	–20.12	–1.92
MP2	–5.80	–16.25
MP3	–17.26	–4.78
MP4SDQ	–10.36	–11.68
MP4SDTQ	–4.51	–17.53
CCSD	–11.06	–10.98
CCSD+T(CCSD)	–6.80	–15.24
CCSDT1	–6.92	–15.12
CCSDT2	–7.56	–14.48
CCSDT3	–7.57	–14.47
6–31++G(2d,p)//SCF/6–31G, vibs. in SCF/6–31++G(2d.p)		
SCF	–18.19	–3.85
MP2	–21.21	–0.83
MP3	–27.82	5.78
MP4SDTQ	–19.37	–2.67
CCD	–23.91	1.87

6–31++G(2d,p)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)

SCF	–24.71	2.67
MP2	–18.15	–3.89
MP3	–29.11	7.07
MP4SDQ	–22.16	0.12
MP4SDTQ	–16.76	–5.28
CCSD	–22.57	0.53
CCSD+T(CCSD)	–19.00	–3.05

6–311G(2df,2pd)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)

SCF	–16.74	–5.30
MP2	–23.17	1.13
MP3	–29.22	7.18
MP4SDTQ	–20.98	–1.06

6–311++G(2df,2pd)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)

SCF	–23.75	1.71
MP2	–21.14	-0.91
MP3	–31.71	9.67
MP4SDQ	–24.33	2.29
MP4SDTQ	-19.20	-2.84

a $\delta\Delta H^{\circ}_{298} = (\Delta H^{\circ}_{298})_{\text{exptl}} - (\Delta H^{\circ}_{298})_{\text{calc}}$, with $(\Delta H^{\circ}_{298})_{\text{exptl}} = -22.04$ kcal and with ΔH°_{298} from E_e values at 0 K without zero-point energies = –20.26 kcal

Table VIII.
Comparison of Various Levels of Approximation for
ΔH°_{298} (in kcal) for the Gaseous Reaction
$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$$

Correlation	ΔH_{298}	$\delta\Delta H^{\bullet}_{298}$
4–31G basis, with vib. approx. in SCF/4–31G		
SCF	–55.57	16.15
MP2(full)	–12.81	–26.61
6–31G basis, with vib. approx. in SCF/6–31G		
SCF	–59.92	20.50
MP2(full)	–17.25	–22.17
MP2	–17.28	–22.14
MP3	–36.34	–3.08
6–31G(d) basis, with vib. approx. in SCF/6–31G		
SCF	–26.14	–13.28
MP2	–7.69	–31.74
MP3	–20.25	–19.18
MP4SDTQ	–3.75	–35.67
CCD	–17.42	–22.00
6–31G(d) basis//Exptl., with vib. approx. in SCF/6–31G		
SCF	–32.56	–6.86
MP2	–5.80	–33.63
MP3	–19.52	–19.90
6–31G(d,p)//SCF/6–31G(d) basis, vibs. in MP2/6–31G(d,p)		
SCF	–36.37	–3.05
MP2	–24.57	–14.85
MP3	–38.21	–1.21
MP4SDQ	–30.99	–8.43
MP4SDTQ	–22.76	–16.66
CCSD	–32.18	–7.24
CCSD+T(CCSD)	–25.91	–13.51
6–31G(d,p)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	–39.13	–0.30
MP2	–21.59	–17.83
MP3	–37.70	–1.72
MP4SDQ	–28.88	–10.54
MP4SDTQ	–18.92	–20.50
CCSD	–30.71	–8.71
CCSD+T(CCSD)	–23.13	–16.29
CCSDT1	–22.84	–16.58
CCSDT2	–24.32	–15.10
CCSDT3	–24.11	–15.31
6–31++G(2d,p)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	–17.81	–21.61
MP2	5.23	–44.66
MP3	–12.50	–26.92
MP4SDQ	–4.17	–35.25
MP4SDTQ	7.75	–47.17
6–311G(2df,2pd)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	–36.77	–2.65
MP2	–37.67	–1.75
MP3	–48.44	9.02
MP4SDTQ	–35.81	–3.(?
6–311++G(2df,2pd)//MP2/6–31G(d,p), vibs. in MP2/6–31G(d,p)		
SCF	–41.88	2.46

a Including only valence excitations (frozen core) except for those denoted as (full).
b $\delta\Delta H^{\circ}_{298} = (\Delta H^{\circ}_{298})_{exptl} - (\Delta H^{\circ}_{298})_{calc}$, with $(\Delta H^{\circ}_{298})_{exptl} = -39.42$ kcal
and ΔH°_{298} from E_e values at 0 K without zero-point energies = –38.49 kcal

Table IX.

Best Values for the Room-Temperature Enthalpy of Formation of Molecules by Dehydrogenation of The Common Hydrides

$$a\ MH_x + b\ LH_z - \tfrac{1}{2}(xa+zb-c)\ H_2 \rightarrow M_aL_bH_c$$

Compound	ΔH°298[a] (kcal/mol)	6-31G[b] Correl.	δΔH	6-31G(d)[c] Correl.	δΔH	6-31G(d,p)// SCF/6-31G(d)[d] Correl.	δΔH	6-31G(d,p)// MP2/6-31G(d,p)[e] Correl.	δΔH	6-31++G(2d,p)// MP2/6-31G(d,p)[f] Correl.	δΔH	6-311G(2df,2pd)// MP2/6-31G(d,p)[g] Correl.	δΔH	Fract. δΔH Vals.[s] <1.1
C2	235.96	MP3	-3.91	CCD	0.79	CCSD	-0.31	CCSD	-0.29	MP4SDQ	-3.77	MP2	3.12	3/45
C2H2	89.95	MP3	8.27	CCD	0.04	CCSD+T[h]	-2.75	CCSD+T	-0.95	MP2	-0.55			3/38
C2H4	48.25	MP2	4.40	MP3	0.28	CCSD+T	0.02	CSDT1	0.25	CSD+T	-0.60			9/43
C2H6	15.52	MP2	0.88	MP3	1.79	MP3	-0.10	MP3	0.10	MP2	-0.82			23/38
CO	49.26	MP3	5.73	SCF	11.77	MP3	1.95	MP3	2.18	CCSD	0.77	MP4[i]	1.80	2/47
CO2	39.42	MP3	3.08	SCF	13.28	MP3	1.21	SCF	0.30			MP2	0.72	2/43
CH2O	47.68	MP3	6.05	SCF	4.59	MP3	-0.23	MP3	0.23	MP4[i]	-0.22			6/38
CH3OH	27.72	SCF	-1.10	SCF	3.45	SCF	1.10	SCF	1.21	MP4	1.16			2/36
(CH3)2O	49.57							SCF	-0.24					1/7
N2	22.04	SCF	-11.17	SCF	12.85	MP3	4.58	SCF	1.92	MP4SDQ	-0.12	MP4	1.06	5/53
N2H4	44.84	MP2	1.95	SCF	2.64	SCF	0.31	SCF	0.35	MP4	-0.83			3/26
HCN	61.20	SCF	-6.65	SCF	1.77	SCF	0.10	MP3	1.40	CCSD	-0.63			4/33
CH3NH2	23.41	MP2	0.28	SCF	-0.15	MP3	0.66							7/15
N2O	99.45	MP3	4.67	SCF	2.50	MP3	-1.41	MP3	-2.37					0/30
HNO3	152.13	MP3	5.25	SCF	5.52									0/11
O3	207.49	MP3	5.62	MP3	12.06	MP4SDQ	-1.96	MP4[h]	2.96					0/33
H2O2	83.01	SCF	3.16	SCF	5.68	SCF	0.03	SCF	-1.77	CSD+T	-1.38			1/37
F2	129.60	SCF	11.45	SCF	6.86	SCF	-0.77	SCF	-3.38	CSD+T	-3.83			1/37
CH2F2	40.68	SCF	2.56	SCF	14.93			SCF	7.50					0/14
CHF3	47.78	SCF	0.34	SCF	25.71									1/6
CF4	56.08	SCF	-4.06	SCF	37.13									0/9
OF2	193.26					SCF	-5.14	SCF	-9.80					0/20
CF2O	53.58	MP3	1.81	SCF	15.83									0/9

[a] The experimental values obtained mainly from Wagman, D. D., et al. *Selected Values of Chemical Thermodynamic Properties*, NBS Tech. Note 270-3, Washington, 1968. with checking against the JANAF Tables for the few of these compounds given therein. |δΔH298| = |(ΔH98)expt. - (ΔH298)theor.|

[b] Correl. limited to (SCF, MP2, and MP3); frequencies in SCF/6-31G.

[c] Correl. limited to (SCF, MP2, MP3, MP4SDTQ, and CCD; frequencies in 6-31G.

[d] Geometry in SCF/6-31G(d); frequencies in MP2/6-31G(d,p). Correl. limited to (SCF, MP2, MP3, MP4SDQ, MP4SDTQ, CCSD, and CCSD+T.

[e] Geometry in MP2/6-31G(d,p); freqs. in MP2/6-31G(d,p). Correl. limited to (SCF, MP2, MP3, MP4SDQ, MP4SDTQ, CCSD, CCSD+T, CCSDT1, and CCSDT2.

[f] Geometry in MP2/6-31G(d,p); frequencies in MP2/6-31G(d,p). Correl. limited to (SCF, MP2, MP3, MP4SDTQ, CCSD, CCSD+T, and CCSD+T.

[g] Gcom. in MP2/6-31G(d,p); freqs. in MP2/6-31G(d,p). Correl. limited to (SCF, MP2, MP3, and MP4SDTQ.

[h] CCSD+T = CCSD+T(CCSD). [i] MP4 =MP4SDTQ.

Table X.

Room–Temperature Enthalpy of Formation of Molecules by Dehydrogenation of The Common Hydrides

$$a\ MH_x + b\ LH_y - \tfrac{1}{2}(xa+zb-c)\ H_2 \rightarrow M_aL_bH_c$$

$\delta\Delta H_{298} = (\Delta H_{298})_{\text{expl.}} - (\Delta H_{298})_{\text{theor.}}$ (kcal/mol of product)

Compound	ΔH°_{298}[a] kcal/mol	6-31G(d)			6-31G(d,p)//SCF/6-31G(d)				6-31++G(2d,p)//MP2/6-31G(d,p)					6-311G-(2df,2pd)		$\Delta\Delta H_{298}$ Dev[c] of Ee vals.
		SCF	MP3	CCD	SCF	MP3	MP4SDQ	CCSD+T[b]	MP2	MP4SDQ	MP4[b]	CCSD	CCSD+T[b]	MP2	MP4[b]	
C_2	235.96	-54.97	-14.49	0.79	-52.72	-23.58	0.80	15.41	4.58	-3.77	15.43	-5.25	11.23	3.13	13.66	2.53
C_2H_2	89.95	-12.78	-1.56	0.04	-9.49	-8.52	-6.37	-2.75	-0.55	-5.75	-1.38	-4.67	-1.33			2.71
C_2H_4	48.25	-7.16	0.28	1.13	-5.14	-3.58	-2.53	0.02	-2.37	-3.36	-1.34	-2.54	-0.60			1.31
C_2H_6	15.52	-3.03	1.79	1.81	-2.65	-0.10	-0.12	0.06	-0.82	-2.29	-1.49	-2.03	-1.39			1.16
CO	49.26	11.77	13.13	16.33	7.29	1.95	7.76	10.36	1.63	1.55	6.01	0.77	4.31	-1.94	1.80	0.93
CO_2	39.41	13.28	19.18	22.00	3.05	1.21	8.43	13.51	(2.65)[d]		(4.10)			0.72	2.59	1.49
CH_2O	47.68	4.59	8.16	9.73	-0.81	0.23	3.17	6.72	-2.28	-3.15	-0.22	-3.27	-0.72			0.37
CH_3OH	27.72	3.45	5.49	5.87	1.10	1.87	2.45	3.52	-1.87	2.06	-1.16	-1.88	-1.16			0.72
N_2	22.04	12.85	16.90	26.11	6.86	4.58	9.60	12.63	0.25	-0.12	5.28	-0.53	3.05	-1.13	1.06	1.78
N_2H_4	44.84	2.64	6.00	6.41	0.31	1.83	2.37	3.41								0.76
HCN	61.20	1.77	10.94	13.39	0.10	0.97	4.81	8.35	4.19	-0.99	4.28	-0.63	3.16			1.93
CH_3NH_2	23.41	-0.15	3.39	3.59	-0.89	0.66	0.94	1.79								
N_2O	99.45	2.49	18.51	22.39												0.75
HNO_3	152.13	5.52	24.42		-10.07	-1.41	8.25	15.53								2.12
O_3	207.49	-20.67	12.06	13.76	-39.32	-9.68	-1.96	10.15	(-12.73)[d]		(-8.33)[d]					-1.27
H_2O_2	83.01	5.68	10.96	11.61	0.03	4.23	5.61	7.58	-5.81	-3.76	-1.65	-3.50	-1.38			0.11
F_2	129.60	6.86	15.65	16.37	-0.77	8.51	10.47	13.15	-10.74	-6.29	-3.75	-5.92	-2.83			0.79
CH_2F_2	40.68	14.92	16.62													0.76
CHF_3	47.78	25.71	29.00													-0.32
CF_4	56.08	37.13	41.69		-5.14	8.70	9.21	18.22								
OF_2	193.26															1.13
CF_2O	53.58	15.83	21.53													0.88

Table XI.
Room–Temperature Enthalpy of Formation of Molecules from Homonuclear Diatomics

$$a/2\ M_2 + b/2\ L_2 + c/2\ H_2 \rightarrow M_a L_b H_c$$

$\delta\Delta H_{298} = (\Delta H_{298})_{expt.} - (\Delta H_{298})_{theor.}$ (kcal/mol of product)

Compound	ΔH°_{298}[a] kcal/mol	6-31G(d)				6-31G(d,p)//SCF/6-31G(d,p)				6-31++G(2d,p)//MP2[b]				6-311G-(2df,2pd)		$\Delta\Delta H_{298}$ Dev[c] of E_e vals.
		SCF	MP2	MP4[b]	CCD	MP4SDQ	MP4	CCSD	CCSD+T[b]	MP4SDQ	MP4	CCSD	CCSD+T[b]	MP2	MP4[b]	
CH_4	-117.98	27.49	-7.35	-13.66	-0.35	-0.40	-9.09	0.16	-7.71	1.89	-7.77	2.64	-5.62	-1.57	-6.88	-2.53
C_2H_2	-146.01	42.19	-12.38	-23.48	-0.75	-7.17	-21.28	-5.08	-18.16	-1.98	-16.81	0.58	-12.06			0.18
C_2H_4	-187.71	47.81	-13.95	-24.32	-0.34	-3.33	-19.03	-1.44	-15.39	0.41	-16.77	2.70	-11.83			-1.22
C_2H_6	-220.44	-51.95	-12.41	-24.81	1.02	-0.92	-17.57	0.37	-14.71	1.47	-16.92	3.21	-12.63			-2.31
NH_3	-22.04	-12.84	-23.73	-26.80	-20.11	-9.60	-13.87	-9.45	-12.63	0.12	-5.28	0.54	-3.05	0.56	-1.06	-0.89
N_2H_4	22.80	-10.20	-17.36	-19.17	-13.70	-7.67	-11.01	-7.31	-9.65							-1.02
HCN	-67.80	22.84	-2.07	-8.02	2.93	-0.17	-6.90	1.12	-5.06	0.96	-6.07	2.26	-3.99			-0.23
CH_3NH_2	-105.59	20.92	-15.20	-22.57	-6.87	-4.90	-14.95	-4.07	-12.87							
HF	-64.80	-3.43	-10.30		-8.19	-5.24	-6.32	-5.37	-6.58	3.14	1.37	2.96	1.42			-0.79
Oxygen Referenced to Water																
O_3	207.49	-20.67	31.13	36.81	13.76	2.96	26.13	2.49	21.34							-1.27
H_2O_2	83.01	5.68	10.79	14.58	11.61	8.54	4.36	8.14	4.24	-3.44	-1.70	-3.40	-1.30			0.11
CO	-68.72	39.26	10.37	9.65	15.94	7.36	2.50	7.23	2.66	3.47	-1.70	3.35	1.30	-3.50	-5.03	-0.10
CO_2	-78.56	40.77	24.38	22.01	21.60	9.90	10.92	8.44	9.84				39.45	-0.85	-4.24	-0.33
CH_2O	-70.30	32.08	4.13	0.82	9.33	3.48	-2.43	3.96	-1.17	-1.27	-7.93	-0.64	-6.34			0.23
CH_3OH	-90.26	30.94	-1.20	-6.46	5.48	2.00	-6.05	2.86	-4.33	-0.17	-8.88	0.74	-6.78			-0.90
N_2O	77.41	-10.35	13.77	15.44	2.27	-0.53	11.62	-3.28	4.73				46.71			-1.04
HNO_3	141.11	-0.91	27.77				19.91						79.24			1.01
OF_2	63.66					4.01	7.79	1.77	3.53							0.35
Oxygen Referenced to Ozone																
H_2O	-69.16	6.89	-10.65	-12.27	-4.59	-0.99	-8.71	-0.83	-7.11				-20.60			0.42
H_2O_2	-55.31	-19.46	-10.51	-9.96	2.44	4.96	-8.22	5.46	-5.46				-42.85			0.96
CO	-137.88	46.15	-0.28	-0.57	11.35	7.54	-4.35	7.31	-2.87				-22.30			0.32
CO_2	-216.88	54.55	3.08	-2.53	12.42	7.92	-6.50	7.18	-5.79				-1.75			0.52
CH_2O	-139.47	38.97	-6.52	-11.46	4.75	2.49	-5.64	3.13	-8.25				-28.53			0.65
CH_3OH	-159.43	-37.83	-11.85	-18.73	0.89	1.02	1.76	2.02	-11.44				-29.48			0.47
N_2O	8.24	-3.46	3.12	3.17	-2.31	-1.51	2.90	-4.11	-2.38				26.10			-0.61
HNO_3	-66.38	19.76	-4.19				-9.21									-0.70
OF_2	-5.51					3.03	-0.93	2.70	0.00							0.77

continuation of Table X.

															35
Oxygen Referenced to Carbon Monoxide															
O_3	413.65	-138.44	0.84	7.87	-34.05	-22.65	13.05	-21.98	8.62		66.91			-9.69	-1.00	-0.97
H_2O	68.72	-39.26	-10.37	-9.65	-15.94	-8.54	-4.36	-8.14	-4.24	-3.44	1.70	-3.35	1.30	3.50	5.03	0.10
H_2O_2	220.45	-72.74	-9.95	-4.72	-20.26	-10.13	0.48	-9.17	0.98	-10.63	1.76	-10.34	1.22	6.16	5.82	0.31
CO_2	58.88	-37.75	3.64	2.71	-10.28	-7.17	2.20	-7.45	0.05		42.85					-0.18
CH_2O	-1.58	-7.18	-6.24	-8.83	-6.60	-5.05	-6.79	-4.19	-5.37	-4.70	-6.23	-4.04	-5.04			0.33
CH_3OH	-21.54	-8.32	-11.57	-16.11	-10.46	-6.53	-10.42	-5.29	-8.57	-3.61	-7.18	-2.66	-5.47			-0.80
N_2O	146.13	-49.61	3.35	5.79	-13.66	-9.06	7.25	-11.42	0.49							-6.93
HNO_3	347.27	-94.08		-9.53			5.66									0.17
OF_2	132.38					-4.83	2.91	-4.94	2.42							0.45
Oxygen Referenced to Carbon Dioxide															
O_3	325.33	-81.82	-4.62	3.80	-18.64	-11.89	9.75	-10.76	8.67		2.63					-0.78
H_2O	39.27	-20.38	-12.19	-11.00	-10.80	-4.95	-5.46	-4.42	-4.22		-19.73			0.42	2.12	0.17
H_2O_2	161.57	-35.09	-13.59	-7.43	-9.99	-2.95	-1.72	-1.72	1.02		-41.20					0.44
CO	-29.44	18.87	-1.87	-1.36	5.14	3.59	-1.10	3.72	0.01		-21.43			-3.08	-2.91	0.06
CH_2O	-30.72	11.99	-7.76	-9.89	-1.17	-1.17	-7.59	-0.16	-5.05		-27.36					0.39
CH_3OH	-50.98	10.52	-13.39	-17.47	-5.32	-2.95	-11.52	-1.56	-8.55		-28.60					0.73
N_2O	116.69	-30.74	1.58	4.44	-8.53	-5.48	5.48	-7.60	-0.51		26.98					0.87
HNO_3	258.95	-62.06		-8.81			2.15									0.14
OF_2	102.94					-0.93	2.32	-0.89	2.90		59.51					0.51

[a] The experimental values were obtained mainly from Wagaman, D. D., et al., *Selected Values of Chemical Thermodynamic Properties*, NBS Tech. Note 270-3, Washington, 1968. with checking against the JANAF Tables for the few of these compounds given therein.

[b] MP4 denotes MP4SDTQ.

Table XII.
Isomerization of Hydrocarbons

Reactant	Product	ΔH°_{298} [b] kcal/mol	Experimental Minus Theoretical Value[a] (kcal/mol)					6–31G(d,p)// * [c]	
			6–31G(d)					MP4 [d]	CCSD+T[d]
			SCF	MP2	MP3	MP4 [d]	CCD		
$H_2C{=}C{=}CH_2$	$HC{\equiv}C{-}CH_3$	−1.24	0.29	2.30	−0.99	−0.26	−0.65	−1.08	−2.20
$H_2C{=}C{=}CH_2$	CH=CH \ / CH$_2$	20.57	−3.60	1.99	−0.12	−0.98	−0.71	−0.56	−1.27
$H_2C{=}CH{-}CH_3$	CH$_2$–CH$_2$ \ / CH$_2$	7.85	−0.89	2.92	1.65	0.56	1.09	0.74	0.12
$H_2C{=}CH{-}HC{=}CH_2$	HC–CH$_2$ ||| H$_2$C–CH	25.79	−4.45	5.38	1.93	0.60	1.04		
$H_2C{=}CH{-}HC{=}CH_2$	CH$_2$–CH$_2$ \ C= / CH$_2$	21.81	1.63	6.01	4.71	4.01	4.40		
trans-CH$_3$CH=CHCH$_3$	cis-CH$_3$CH=CHCH$_3$	1.13	−0.63	−0.58	−0.55	−0.52	−0.63		
trans-CH$_3$CH=CHCH$_3$	H$_2$C–CH$_2$ | | H$_2$C–CH$_2$	9.77	0.49	3.91	3.14	2.29	2.88		

[a] $\delta\Delta H^\circ_{298}$

[b] The experimental values were obtained from J. D. Cox and G. Pitcher, *Thermochemistry of Organic and Organometallic Compounds*, Academic, London, 1970.

[c] 6–31G(d,p)// * stands for 6–31G(d,p)//SCF/6–31G(d)

[d] MP4 stands for MP4SDTQ, and CCSD+T for CCSD+T(CCSD)

Table XIII
Enthalpies (kcal/mol at 25°C) of Formation of Hydrocarbons
from Methane Dehydrogenation: $nCH_4 - (2n - x/2)H_2 \rightarrow C_nH_x$

Compound	Exptl. Value[b] (kcal/mol)	Experimental Minus Computed Value[a] (kcal/mol)									
		6-31G(d)				6-31G(d,p)//SCF/6-31G(d)					
		MP2	MP3	MP4[c]	CCD	MP2	MP3	P4-T[c]	MP4[c]	CCSD	SD+T[c]
C_2H_2	89.95	2.33	-1.56	3.84	0.04	-3.49	-9.04	-6.89	-3.62	-5.91	-3.27
C_2H_4	48.25	0.76	0.28	3.01	1.13	-2.23	-3.68	-2.63	-0.96	-1.86	0.08
C_2H_6	15.50	2.30	1.79	2.51	1.81	1.07	0.10	0.08	0.80	0.26	0.87
$H_2C=C=CH_2$	99.27	3.43	1.57	7.54	2.93	-3.15	-7.09	-5.05	-1.18	-3.83	0.01
$HC\equiv CCH_3$	98.03	5.73	0.57	7.28	2.28	-1.64	-8.89	-6.44	-2.25	-5.52	-2.19
CH=CH \ / CH₂	119.84	5.42	1.44	6.57	2.23	-0.83	-8.85	-5.88	-1.74	-4.81	-1.26
$H_2C=CHCH_3$	58.52	3.59	2.07	5.77	2.75	-0.94	-3.97	-2.99	-0.31	-2.16	0.42
CH₂-CH₂ \ / CH₂	66.37	6.51	3.72	6.33	3.84	2.10	-2.19	-2.09	0.43	-1.50	0.54
$H_3CCH_2CH_3$	28.81	4.99	3.66	5.33	3.63	2.39	0.09	0.43	1.75	0.37	1.78
$HC\equiv CC\equiv CH$	184.52	13.47	1.77	15.67	5.06	0.62	-14.59	-9.69	-0.70	-8.19	-1.34
CH₂-CH₂ \ / C ‖ CH₂	119.44	10.64	6.20	12.51	7.06						
$H_2C=CH\text{-}HC=CH_2$	97.63	4.63	1.49	8.50	2.66						
HC-CH₂ I\I H₂C-CH	123.42	10.01	3.42	9.10	3.70						
cis- $H_3CHC=CHCH_3$	69.66	5.55	2.94	7.74	3.39						
trans- $H_3CHC=CHCH_3$	68.53	6.14	3.49	8.25	4.02						
$CH_3CH_2CH_2CH_3$	41.12	6.78	4.58	7.25	4.48						

[a] $\delta\Delta H^o_{298}$. Vibrational corrections computed in the 6-31G basis. [b] Experimental ΔH^o_{298} for the reaction.
[c] MP4 stands for MP4SDTQ, P4-T for P4SDQ, and SD+T for CCSD+T(CCSD).

Table XIV.
Enthalpies (kcal/mol at 25°C) of Formation of Hydrocarbons
from Hydrogenation of Diatomic Carbon: $n/2 C_2 + x/2 H_2 \rightarrow C_n H_x$

Compound	Exptl. Value[b] (kcal/mol)	Experimental Minus Computed Value[a] (kcal/mol)									
		6–31G(d)				6–31G(d,p)//SCF/6–31G(d)					
		MP2	MP3	MP4[c]	CCD	MP2	MP3	P4–T[c]	MP4[c]	CCSD	SD+T[c]
CH_4	−117.98	−7.35	7.25	−13.66	−0.40	−4.17	11.37	−0.82	−9.51	−0.26	−8.13
C_2H_2	−146.01	−12.38	12.93	−23.48	−0.75	−11.83	13.71	−8.53	−22.64	−6.44	−19.52
C_2H_4	−187.71	−13.95	14.77	−24.32	0.34	−10.57	19.06	−4.27	−19.98	−2.38	−16.33
C_2H_6	−220.44	−12.41	16.28	−24.82	1.02	−7.27	22.84	−1.56	−18.21	−0.27	−15.38
$H_2C=C=CH_2$	−254.67	−18.63	23.30	−33.45	1.74	−15.66	27.02	−7.52	−29.70	−4.62	−24.37
$HC≡CCH_3$	−255.91	−16.34	22.31	−33.71	1.10	−14.16	25.22	−8.91	−30.77	−6.31	−26.57
CH=CH \ / CH$_2$	−234.10	−16.64	23.18	−34.42	1.04	−13.34	27.26	−8.39	−30.26	−5.60	−25.64
$H_2C=CH–CH_3$	−295.42	−18.47	23.80	−35.22	1.57	−13.45	30.14	−5.45	−28.89	−2.95	−23.96
CH$_2$–CH$_2$ \ / CH$_2$	−287.57	−15.55	25.45	−34.66	−2.65	−10.41	31.92	−4.55	−28.10	−2.29	−23.84
$H_3CCH_2CH_3$	−325.13	−17.07	25.40	−35.66	2.44	−10.12	34.20	−2.42	−26.81	−0.42	−22.60
$HC≡CC≡CH$	−287.40	−15.95	30.75	−38.99	3.48	−16.06	30.89	−12.98	−38.73	−9.24	−33.84
CH$_2$–CH$_2$ \ / C ‖ CH$_2$	−352.48	−18.78	35.18	−42.14	5.48						
$H_2C=CH{-}HC=CH_2$	−374.29	−24.79	30.47	−46.15	1.08	−20.22	35.39	−10.09	−37.65		
HC–CH$_2$ I\I H$_2$C–CH	−348.50	−19.41	32.41	−45.55	2.12						
cis-$H_3CHC=CHCH_3$	−402.26	−23.86	31.93	−46.92	1.81						
trans-$H_3CHC=CHCH_3$	−403.39	−23.28	32.48	−46.40	2.44						
$CH_3CH_2CH_2CH_3$	−430.80	−22.64	33.57	−47.40	2.90						

[a] $\delta \Delta H^o_{298}$. Vibrational corrections computed in the 6-31G basis. [b] Experimental ΔH^o_{298} for the reaction.
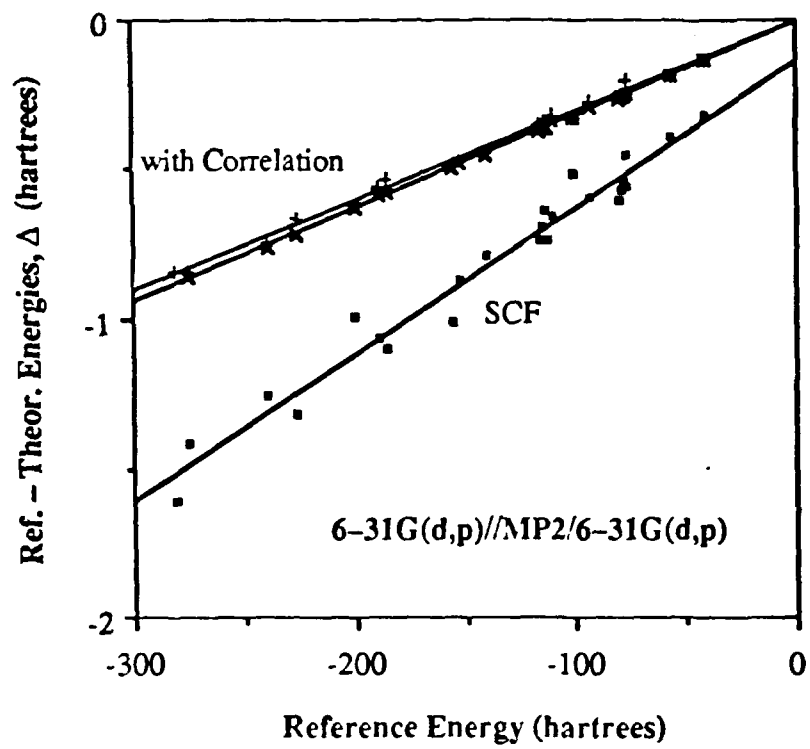[c] MP4 stands for MP4SDTQ, P4–T for P4SDQ, and SD+T for CCSD+T(CCSD).

Figure Titles

Figure 1. Plots, for a series of molecules without nuclear motion of the disagreement between the reference and theoretical energies, $\Delta = E_e - E_{theor}$, as a function of the reference energy, E_e, for SCF computations without (•) and with MP4SDTQ (×) and CCSD (⁺) correlations in the 6–31G(d,p)//MP2/6–31G(d,p) description. Some points for the two correlation schemes are superimposed at the scale of this figure.

Figure 2. Plots for a series of molecules without nuclear motion of the disagreement between the reference and theoretical energies, $\Delta = E_e - E_{theor}$, as a function of the reference energy, E_e, for SCF computations in the small 6–31G//SCF/6–31G (•) and the large 6–311++G(2df,2pd)//MP2/6–31G(d,p) (×) description.

Figure 3. Variations for a series of molecules without nuclear motion of the intercept-corrected percentage disagreement between the reference and theoretical energies, $\%(\Delta-\alpha) = 100 (E_e - E_{theor} - \alpha)/E_e$, vs. the reference energy, E_e, in two plots representing (above) the 6–31++G(d,p)//MP2/6–31G(d,p) and (below) the 6–31G(d,p)//MP2/6–31G(d,p) description. In each plot, the SCF is shown for $\alpha = -0.138$ {%Δ SCF} and for $\alpha = 0$ {%(Δ–0.138)}, below which appears the data for $\%(\Delta-\alpha)$ with $\alpha = 0$ for CCSD correlation {%Δ CCSD} and for MP4SDTQ correlation {%Δ MP4}. The horizontal lines projecting from the vertical axis correspond to average values. Data from Table I.

Figure 4. Relationship between the percentage difference, %Δ, in the 6–31G(d) basis set and the H/C atomic ratio for some typical hydrocarbons without nuclear motion.

Figure 5. Relationships between the percentage difference, %Δ, in the 6–31G(d,p)//MP-/6-31G(d,p) description and the H/X atomic ratio for some molecules containing two different kinds of first-row atoms, where the symbol X stands for the atoms involved and their ratios. Thus, X=(C,2F) stands for all of those molecules having a carbon:fluorine ratio of 1:2 with the H/X axis giving their individual hydrogen:(carbon + 2xfluorine) atomic ratio.

Figure 6. Comparison of $\delta\Delta H^\circ_{298}$ for four formation reactions for a series of hydrocarbons studied in the 6–31G* basis and shown as separate graphs for each correlation scheme: MP3, MP4SDTQ, and CCD, reading from top to bottom. The plot for MP2 was not depicted because it is quite similar to that for MP4SDTQ. The reactions deal with formation from **A**, dehydrogenation of CH_4; **B**, hydrogenation of C_2; **C**, hydrogenation of CO with dehydration, and **D**, hydrogenation of CO_2 with dehydration.
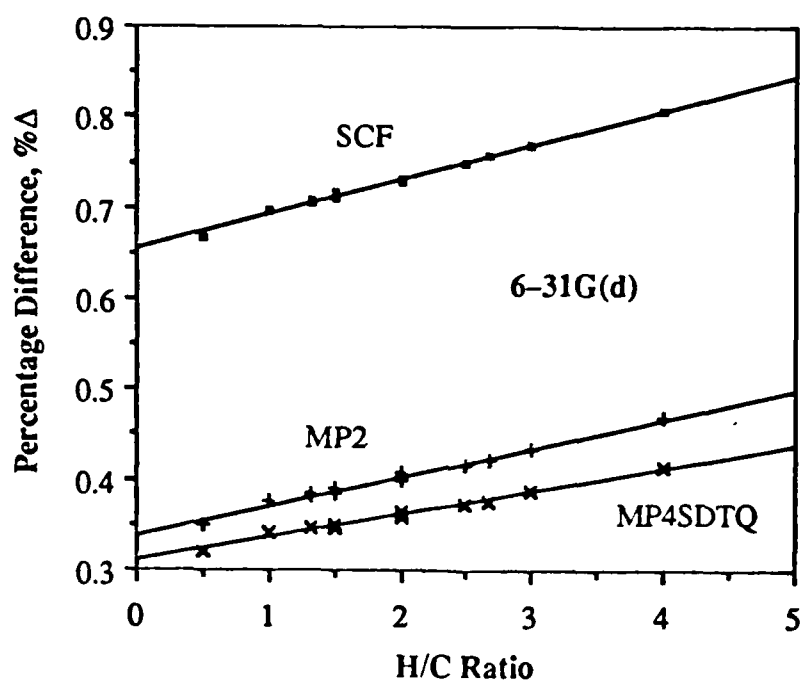
Fig. 1

Fig 2

Fig 3

Fig 4

6–31G(d,p)//MP2/6–31G(d)

Fig 5

Fig 6